

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
14 December 2000 (14.12.2000)

PCT

(10) International Publication Number
WO 00/75744 A2

(51) International Patent Classification⁷: G06F
(21) International Application Number: PCT/IB00/00853
(22) International Filing Date: 7 June 2000 (07.06.2000)
(25) Filing Language: English
(26) Publication Language: English
(30) Priority Data:
09/327,049 7 June 1999 (07.06.1999) US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant: NORTEL NETWORKS LIMITED
[CA/CA]; World Trade Center of Montreal, 8th floor,
380 St. Antoine Street West, Montreal, Quebec H2Y 3Y4
(CA).

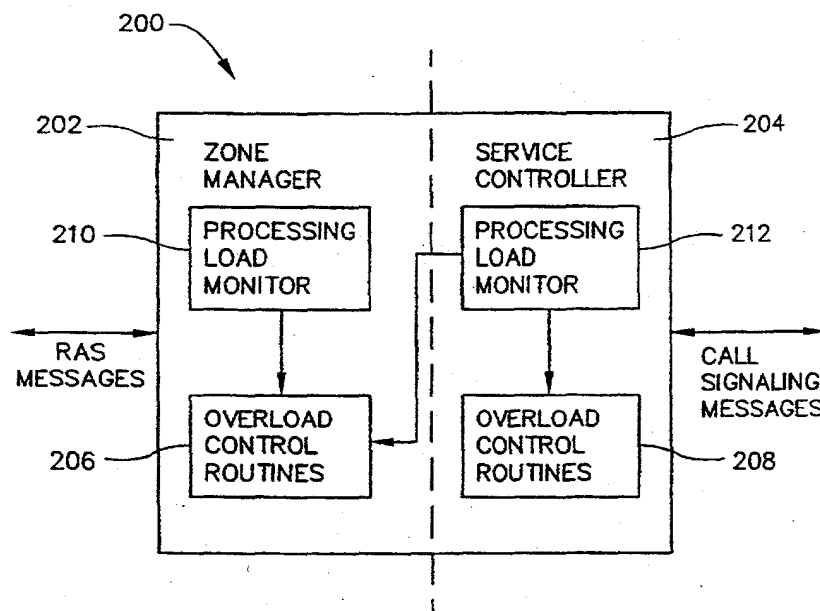
Published:

— Without international search report and to be republished upon receipt of that report.

(72) Inventors: LI, Xuewen; 105 Modena Drive, Cary, NC 27513 (US). CHRISTIE, Samuel, H., IV; 309 Trapper's Run Drive, Cary, NC 27513 (US).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS AND SYSTEMS FOR CONTROLLING NETWORK GATEKEEPER MESSAGE PROCESSING



(57) Abstract: Methods and systems for controlling network gatekeeper message processing include overload control routines executed by a network gatekeeper to prevent the network gatekeeper from crashing during overload conditions. The overload control routines are executed in a predetermined order so that messages relating to calls in progress are preferentially processed over messages relating to newly originating calls. Fault detection routines detect protocol stack and application level faults of a network

WO 00/75744 A2

-1-

DescriptionMETHODS AND SYSTEMS FOR CONTROLLING NETWORK GATEKEEPER
MESSAGE PROCESSINGTechnical Field

5 The present invention relates generally to methods and systems for controlling a network gatekeeper. More particularly, the present invention relates to methods and systems for controlling message processing by an H.323 gatekeeper.

Related Art

10

H.323 is a set of recommendations from the International Telecommunications Union (ITU) for multimedia communications over local area networks (LANs) that do not provide guaranteed quality of service (QoS). These
15 networks are commonly used in corporate and educational institutions. Exemplary packet-switched networks that do not guarantee quality of service include TCP/IP, UDP/IP, and IPX over Ethernet, Fast Ethernet and Token Ring. H.323 provides recommendations for communication between
20 voice-only handsets, video conferencing stations, and other devices capable of transmitting voice, video, and data packets over such networks.

H.323 defines four major components for a network-based communication system: terminals, gateways,
25 gatekeepers, and multi-point control units. Terminals are client endpoints connected to a LAN that provide real-time, two-way communications. A terminal may include a telephone capable of transmitting voice packets over a network, a video conferencing station capable of
30 transmitting voice and video over a network, or a computer

-2-

capable of transmitting voice, video, and data over a network. A gateway is an optional element of an H.323 network that provides translation between transmission formats and between communications procedures. For example, a gateway may establish links between H.323 terminals and analog public switched telephone network (PSTN) terminals. Thus, gateways function as translators between H.323 compliant devices and non-H.323 compliant devices.

10 A gatekeeper is the most important part of an H.323 enabled network. An H.323 gatekeeper functions as the central point for all calls within its zone and provides call control services to the registered endpoints. Figure 1 illustrates a conventional H.323 gatekeeper and an H.323 zone. In Figure 1, an H.323 gatekeeper 100 manages message traffic in an H.323 zone 101. An H.323 zone is defined by the collection of terminals, gateways, and multi-point control units managed by a single gatekeeper. In the illustrated example, the H.323 zone 101 spans multiple networks 102 and 103 connected by routers 104 and 105. The H.323 zone 101 includes terminals 106-110 that transmit packets containing multi-media data to each other in accordance with H.323 recommendations. A gateway 111 translates between transmission formats and communications procedures. A multi-point control unit (MCU) 112 supports conferences between three or more endpoints.

Gatekeepers perform two important call control functions. The first function is address translation from LAN aliases for terminals and gateways to IP or IPX addresses, as defined in the Registration, Admission, Status (RAS) specifications included in ITU Recommendation H.225. The second function is bandwidth management, which is also described in ITU Recommendation H.225. For

-3-

example, a network manager may have a specific threshold for the number of simultaneous conferences on a local area network (LAN). The gatekeeper can refuse to make connections once the threshold is reached. The effect of
5 refusing to make connections is to limit the total conferencing bandwidth to some fraction of the total available network bandwidth. The remaining bandwidth may be used for e-mail, file transfers, and other LAN protocols.

10 An optional, but valuable, feature of a gatekeeper is the ability to route H.323 calls. By routing a call through a gatekeeper, the call can be controlled more effectively. Service providers need this ability in order to bill for calls placed through their networks. This
15 service can also be used to re-route a call to another endpoint if a called endpoint is unavailable. In addition, a gatekeeper capable of routing H.323 calls can assist in making decisions involving balancing among multiple gateways. For instance, if a call is re-routed
20 through a gatekeeper, that gatekeeper can then re-route the call to one of the many gateways based on a balancing algorithm. While a gatekeeper is logically separate from H.323 endpoints, vendors may incorporate gatekeeper functionality into physical implementation of gateways and
25 MCUs.

A gatekeeper is not required in an H.323 system. However, if a gatekeeper is present, terminals must make use of the services offered by the gatekeepers. RAS, as described in ITU Recommendation H.225, defines these
30 services as address translation, admissions control, bandwidth control, and zone management.

Gatekeepers can also play a role in multi-point communications. To support multi-point conferences, users

may employ a gatekeeper to receive control channels from two terminals in a point-to-point conference, as described in ITU Recommendation H.245. When the conference switches to multi-point, the gatekeeper can direct the H.245 control channel to a multi-point controller, or MC. The gatekeeper need not process the H.245 signaling; it only needs to pass the H.245 signals between the terminals or between the terminals and the MC.

LANs containing gateways can also contain a gatekeeper to translate incoming E.164 addresses to transport addresses. Because a zone is defined by its gatekeeper, H.323 entities that contain an internal gatekeeper require a mechanism to disable the internal function so that if there are multiple H.323 entities that contain a gatekeeper on a LAN, the entities can be configured into the same zone. Tables 1 and 2 shown below illustrate required and optional gatekeeper functions, respectively.

Table 1: Required Gatekeeper Functions

20

Address Translation	Translation of an alias address to a transport address using a table that is updated with registration messages. Other methods of updating the translation table are also allowed.
----------------------------	--

-5-

Admissions Control	Authorization of LAN access using Admissions Request, Confirm and Reject (ARQ/ARC/ARJ) messages. LAN access can be based on call authorization, bandwidth, or some other criteria. Admissions Control can also be a null function that admits all requests.
Bandwidth Control	Support for Bandwidth Request, Confirm and Reject (BRQ/BCF/BRJ) messages. This can be based on bandwidth management. Bandwidth Control can also be a null function which accepts all requests for bandwidth changes.
Zone Management	The gatekeeper provides the above functions for terminals, MCUs, and gateways which have registered within its zone of control.

Table 2: Optional Gatekeeper Functions

5

Call Control Signaling	In a point-to-point conference, the gatekeeper can process Q.931 call control signals. Alternatively, the gatekeeper can send the endpoints' G.931 signals directly to each other.
Call Authorization	The gatekeeper can reject a call from a terminal based on the Q.931 specification. The reasons for rejection can include, but are not limited to, restricted access to/from particular terminals or gateways, restricted access during certain periods of time. The criteria for determining if authorization passes or fails is outside the scope of H.323.

Bandwidth Management	The gatekeeper can reject calls from a terminal if it determines that sufficient bandwidth is not available. This function also operates during an active call if a terminal requests additional bandwidth. The criteria for determining if bandwidth is available is outside the scope of H.323.
Call Management	The gatekeeper can maintain a list of ongoing H.323 calls in order to indicate that a called terminal is busy or to provide information for the Bandwidth Management function.

Because gatekeepers execute on one or more physical machines, the processing ability of a gatekeeper is limited. In light of all of the important functions performed by a gatekeeper, it is preferable to define methods for controlling gatekeeper message processing and for preventing the gatekeeper from crashing during overload conditions. As used herein, the phrases "overload conditions" and "overloaded state" refer to the consumption of a percentage of gatekeeper processing capacity. For example, if the processing capability of the machine on which the gatekeeper executes is seventy percent consumed, the gatekeeper might be determined to be in an overloaded state.

One conventional method for preventing a gatekeeper from crashing during overload conditions is to cease accepting incoming messages when the gatekeeper is overloaded. One problem with this approach is that there is no mechanism for favoring calls in progress over newly originating calls. As a result, calls that are almost complete are abandoned with equal frequency as newly

-7-

originating calls. The equal frequency of abandonment wastes network resources and results in loss of revenue by the service provider.

Another conventional approach for managing a network gatekeeper in overload conditions is to discard messages received by a gatekeeper based on message type. One problem with this approach is that the message type alone does not determine the semantics of the message. For example, a Setup message normally corresponds to a call origination. However, if emergency 911 is active, the Setup message can be a reconnect to the emergency operator, as opposed to a new call request. A management program that identifies and discards Setup messages based on message type only can thus discard emergency 911 reconnect Setup messages in addition to new origination Setup messages. Discarding emergency 911 reconnect Setup messages is undesirable. Because the message type alone might not provide sufficient information to determine whether to discard an incoming message, discarding calls based solely on message type might not be an effective way to regulate gatekeeper message processing during overload conditions.

Accordingly, there exists a long felt need for methods and systems for controlling gatekeeper message processing during overload conditions.

Disclosure of the Invention

The present invention provides methods and systems for controlling gatekeeper message processing, especially during overload conditions. In order to reduce the likelihood of gatekeeper failure in overload conditions, lines of defense can be provided for regulating incoming message traffic. One line of defense includes discarding

incoming messages in a manner structured to favor calls in progress over newly originating calls. Another line of defense includes prioritizing calls based on user profile or called party destination address. Prioritizing incoming calls can include preferentially allocating gatekeeper processing resources to calls of higher priority. Another line of defense can include detecting and discarding abandoned calls.

According to another aspect, the present invention can include methods and systems for detecting gatekeeper faults. For example, in order to determine whether a gatekeeper's network protocol stack is functioning properly, a ping message can be transmitted to the gatekeeper. In order to determine whether the gatekeeper is capable of receiving calls, a Gatekeeper Confirm (GCF) message can be transmitted to the gatekeeper.

Accordingly, it is therefore an object of the present invention to provide novel methods and systems for regulating gatekeeper message processing.

Brief Description of the Drawings

A description of the present invention will now proceed with reference to the accompanying figures of drawings of which:

Figure 1 is a block diagram illustrating a gatekeeper and conventional H.323 zone;

Figure 2 is a block diagram of a gatekeeper including a zone manager and a service controller having overload control routines according to an embodiment of the present invention;

Figure 3 is a flow diagram illustrating message flow between a calling party and a gatekeeper for a typical H.323 call;

Figure 4 is a flow chart illustrating an exemplary hierarchy for overload control routines according to an embodiment of the present invention;

5 Figure 5 is a flow chart illustrating a user profile prioritizer according to an embodiment of the present invention;

Figure 6 is a flow chart illustrating an input handler according to an embodiment of the present invention;

10 Figure 7 is a flow diagram illustrating origination and termination ARQ messages;

Figure 8 is a flow chart illustrating a zone manager admissions controller according to an embodiment of the present invention;

15 Figure 9 is a flow chart illustrating a called party prioritizer according to an embodiment of the present invention;

Figure 10 is a flow chart illustrating a service controller origination limiter according to an embodiment
20 of the present invention;

Figure 11 is a flow chart illustrating a call processor with in-progress message favoring according to an embodiment of the present invention;

Figure 12 is a flow chart illustrating an abandoned
25 call detector according to an embodiment of the present invention; and

Figure 13 is a flow chart illustrating a zone manager fault detector according to an embodiment of the present invention.

30 Detailed Description of the Invention

Figure 2 illustrates a network gatekeeper including overload control routines according to an embodiment of

-10-

the present invention. In Figure 2, a network gatekeeper generally designated 200 can comprise one or more processes executing on a single machine or a plurality of machines. For example, the gatekeeper can execute on one or more personal computers. Gatekeeper 200 can be capable of providing multi-media communication services over a packet-based network that does not provide guaranteed quality of service. In a preferred embodiment, gatekeeper 200 comprises an H.323 gatekeeper. In Figure 2, gatekeeper 200 includes a zone manager 202 and a service controller 204. Zone manager 202 comprises a process or collection of processes within gatekeeper 200 that perform RAS functions within a zone. Service controller 204 comprises a process or collection of processes within gatekeeper 200 that perform call signaling functions. Specific functions performed by zone manager 202 and the service controller 204 will be described in more detail below.

According to a preferred embodiment of the present invention, zone manager 202 includes one or more overload control routines 206 for controlling zone manager message processing when gatekeeper 200 is in an overloaded condition. Overload control routines 206 can be variously configured. For example, overload control routines 206 can include routines for prioritizing incoming calls based on a user profile or a called party. Overload control routines 206 can also include routines that favor calls in progress over newly originating calls. The specific structure of overload control routines 206 will be discussed in more detail below.

In order to determine whether overload conditions exist at the zone manager or the service controller, both the zone manager and the service controller preferably

-11-

include processing load monitors. For example, in the illustrated embodiment, zone manager 202 includes a processing load monitor 210 that determines whether the zone manager is in an overloaded condition and the extent of the processing load on the zone manager. Service controller 204 preferably also includes a processing load monitor 212 to determine whether the service controller is in an overloaded condition and the extent of the service controller processing load. Processing load monitors 210 and 212 can use any indicator to determine whether zone manager 202 and service controller 204 are overloaded and the extent of the processing loads. For example, both processing load monitors 210 and 212 can measure the percentage of CPU capacity being consumed or the remaining CPU capacity in the machine(s) executing zone manager 202 and service controller 204. Based on the percentages measured by processing load monitors 210 and 212, zone manager 202 and service controller 204 can execute one or more of overload control routines 206 and 208.

Zone Manager/Service Controller Communications

Because zone manager 202 and service controller 204 can be separate processes that can execute on separate machines, it is preferable that service controller 204 communicate its available processing capacity to zone manager 202. Any manner of communicating available processing capacity from service controller 204 to zone manager 202 is within the scope of the invention. For example, if zone manager 202 and service controller 204 have access to the same memory, processing load monitor 212 of service controller 204 can set a flag that indicates available or consumed service controller

-12-

processing capacity. Zone manager 202 can periodically check the flag to determine the available or consumed processing capacity of service controller 204. Alternatively, processing load monitor 212 of service controller 204 can transmit a message to zone manager 202 to communicate available service controller processing capacity to zone manager 202. Communicating available or consumed service controller processing capacity to zone manager 202 enables zone manager 202 to implement initial lines of defense against gatekeeper failure based on the processing load on service controller 204.

The overload control routines, such as overload control routines 206 and 208, according to the present invention regulate gatekeeper message processing based on a number of criteria, including message content and relationships between messages in a call. Thus, a typical H.323 call sequence will now be described as Figure 3 illustrates a typical call sequence between an H.323 client and a gatekeeper. The call sequence in Figure 3 illustrates gatekeeper routed signaling, which is the type of signaling to which the overload control routines according to the present invention are preferably applied.

In order to initiate a call on an H.323 network, a client, such as a telephone, a computer, or a video conferencing terminal, must discover the network gatekeeper and register with the network gatekeeper. In order to discover the network gatekeeper, the client transmits a Gatekeeper Request (GRQ) message to the network gatekeeper. The zone manager of the network gatekeeper responds with a Gatekeeper Confirm message (GCF) indicating the transport address of the gatekeeper RAS channel. After identifying the gatekeeper, in order

-13-

to register with the gatekeeper, the client sends a Registration Request (RRQ) to the gatekeeper through the RAS channel. The zone manager of the gatekeeper responds with a Registration Confirm (RCF) indicating that the client is registered and can proceed to make a call. The client then transmits an Admissions Request (ARQ) message that asks the zone manager for permission to make a call, specifies the bandwidth for the call, and the destination of the call. If the zone manager accepts the call, the zone manager responds with an Admissions Confirm (ACF) message. The Admissions Confirm message includes the IP address of the service controller portion of the gatekeeper.

An exemplary implementation of a gatekeeper separated into a zone manager and a service controller is described in commonly-assigned, co-pending U.S. patent application number 09/249,051, filed February 12, 1999 and entitled, "System and Method for Providing Service Control to a Single Telephone End Terminal from Multiple Service Providers", the disclosure of which is incorporated herein by reference in its entirety.

In response to receiving the Admissions Confirm message, the client sends a Setup message to the service controller. The service controller responds by sending a Call Proceeding message to the client. The Call Proceeding message informs the client that the service controller is searching for the called party address. Once the service controller locates the called party, the service controller presents the called party with the call and transmits an Alerting message to the client. The Alerting message informs the client that the called party has been presented with the call. If the called party is a telephone, the Alerting message indicates that the

-14-

called party's telephone is ringing. When the called party answers the call, the service controller transmits a Connect message to the client. Once the client receives the Connect message, the client can transmit voice, video, and/or data to the called party over the H.323 network. When the called party goes on-hook, the service controller sends a Release Complete message to the client. In response to receiving the Release Complete message, the client sends a Disengage Request (DRQ) message to the service controller. The DRQ message instructs the service controller to deallocate any bandwidth previously allocated for the call.

Overload Control Routines

As stated above, both the zone manager and the service controller preferably include one or more overload control routines. The overload control routines are preferably executed in a predetermined hierarchy based on available or consumed zone manager and service controller processing capacity. For example, the zone manager processes RAS messages, and the service controller processes call signaling messages. Since RAS messages occur before call signaling messages in a call, fewer gatekeeper resources have been expended on a call at the RAS message processing stage than at the call signaling processing stage. Thus, it is preferable to execute zone manager overload control routines before service controller overload control routines. In addition, hierarchies preferably exist within the zone manager and the service controller overload control routines that favor the processing of messages relating to calls for which more resources have been expended.

-15-

Figure 4 illustrates an exemplary hierarchy for overload control routines according to an embodiment of the present invention. In Figure 4, the routines above the dashed line are included in the zone manager, and the routines below the dashed line are included in the service controller. A user profile prioritizer 400 is a routine included in the zone manager that bypasses the overload control routines when a call is from a priority user. Similarly, a called prioritizer 402 bypasses overload control routines below the called party prioritizer 402 in Figure 4 when the call is to a priority destination, such as an emergency 911 operator. The remaining routines in Figure 4 represent a series of gates through which messages must pass in order to be processed by the gatekeeper. One or more of the routines may be simultaneously active in order to reduce the likelihood of gatekeeper failure.

User Profile Prioritizer

User profile prioritizer 400 is included in the zone manager to prevent messages relating to calls from priority users from being discarded. For example, an executive employed by a corporation having an H.323 network can be considered a high-priority user. As a result, calls originating from the corporate executive's telephone can be given priority over calls originating from a lower-level employee's telephone. Resources in the zone manager can be reserved in advance for processing calls from priority users.

Figure 5 illustrates an exemplary user profile prioritizer that may be included in the zone manager of a network gatekeeper. In step ST1, the user profile

-16-

prioritizer receives an incoming call Admissions Request. The incoming call Admissions Request can include the IP address of the call originator. In step ST2, the user profile prioritizer extracts the IP address of the originator. In steps ST3 and ST4, the user profile prioritizer determines whether the originator is a priority user. For example, the user profile prioritizer can compare the IP address of the originator to a plurality of IP addresses stored in a lookup table to determine a profile for the call originator. If the profile indicates that the user is a priority user, such as a corporate executive, the user profile prioritizer can process the message using reserved resources (step ST5) and bypass the remaining overload control routines illustrated in Figure 4. Exemplary resources that can be reserved include gatekeeper processor time and memory needed to process a call. If the IP address does not indicate that the originator is a priority user, the call can be processed normally (step ST6), i.e., without bypassing the overload control routines. The user profile prioritizer ensures that messages from priority users are processed even when the gatekeeper enters an overloaded condition.

Input Handler

Since RAS messages are the initial messages in a call, and the zone manager handles RAS messages, a first line of defense in gatekeeper overload conditions may be to limit incoming RAS messages. Referring back to Figure 4, the zone manager may include an input handler 404 that preferentially discards RAS messages based on the amount of processing resources expended on the call to which the RAS messages pertain. For example, input handler 404 can

preferentially discard RAS messages relating to newly originating calls over RAS messages relating to calls in progress.

Figure 6 illustrates exemplary steps that can be performed by input handler 404, according to an embodiment of the present invention. The steps illustrated in Figure 6 can be executed when zone manager or service controller processing load reaches a first level, for example, ninety percent of total processing capacity. When this threshold is reached, the input handler can preferentially discard ARQ messages over DRQ messages. Referring back to Figure 3, an ARQ message can represent a new origination and a DRQ message is associated with a call in progress. ARQ messages can be distinguished from DRQ message by examining the message type field of an incoming RAS message. Thus, identifying and distinguishing between ARQ and DRQ messages requires minimal processing by the input handler.

In step ST1 of Figure 6, the input handler can discard incoming ARQ messages in order to reduce zone manager or service controller processing load. If processing load decreases below the level that triggered the input handler, the input handler can stop discarding ARQ messages (steps ST2 and ST3). If discarding incoming ARQ messages is not sufficient to reduce processor load below the triggering threshold, the input handler can discard incoming DRQ messages (step ST4).

Zone Manager Admissions Controller

A second line of defense that can be implemented with regard to ARQ messages is a zone manager admissions controller that preferentially discards origination ARQ

messages over termination ARQ messages. Figure 7 is a flow diagram illustrating the difference between origination and termination ARQs. In line 1), a first terminal, terminal A, in an H.323 zone transmits a Setup message to a gatekeeper directing the gatekeeper to set up a call with a second terminal, terminal B. The Setup message includes a call reference value of X. The call reference value is a number included in messages that identifies the call to which the messages pertain. In line 2), the gatekeeper responds by transmitting a Setup message to terminal B with the call reference value of X. In line 3), terminal B responds to the Setup by transmitting an ARQ to the gatekeeper with the call reference value of X. The ARQ message illustrated in line 3) is referred to as a termination ARQ because it is transmitted by a called party. Since the zone manager receives ARQ messages and the service controller receives Setup messages, identifying termination ARQ messages can require communication between the zone manager and the service controller of the gatekeeper.

The dashed line in Figure 7 represents a division between the first call and a second call. In line 4), terminal B transmits an ARQ message to the gatekeeper for a new call. The ARQ has a call reference value of Y. The ARQ message illustrated in line 4) is an origination ARQ because it is transmitted by a calling party to establish a new call. Because more call processing resources have been expended on a termination ARQ than an origination ARQ, the gatekeeper admissions limiter preferably discards origination ARQs before discarding termination ARQs.

As illustrated in Figure 4, zone manager admissions controller 406 can be executed when the processor load of the zone manager and/or the service controller reaches a

-19-

second state, higher than the state required to trigger execution of the input handler illustrated in Figure 6. Figure 8 illustrates exemplary steps that can be executed by a zone manager admissions controller according to an embodiment of the present invention. In step ST1, the zone manager admissions controller identifies and discards origination ARQ messages. Identifying ARQ messages can include examining the message type to determine whether the message is an ARQ and examining the call reference value to determine whether the message relates to a new call. If a message is an ARQ for a new call, the zone manager admissions controller preferably discards the ARQ message. In step ST2, the zone manager admissions controller determines whether the processing load of the zone manager or the service controller has sufficiently decreased, e.g., whether the processing load is below the threshold that triggered the zone manager admissions controller. If discarding originating ARQs causes the percentage of available processing capacity to increase, the zone manager admissions controller can stop discarding origination ARQs (step ST3). If the processing capacity is not sufficiently increased or, alternatively, if the processing load is not sufficiently decreased, then the zone manager admissions controller preferably identifies and discards termination ARQs, in addition to origination ARQs (step ST4). By favoring the processing of termination ARQs over origination ARQs, the zone manager admissions controller efficiently regulates gatekeeper message processing at the zone manager level.

30

Called Party Prioritizer

Referring back to Figure 4, the next level of message processing regulation is called party prioritizer 402. Called party prioritizer 402 can be a component of the service controller of the gatekeeper that prevents messages addressed to a priority destination, such as an emergency 911 operator, from being discarded. Figure 9 illustrates exemplary steps that can be included in a called party prioritizer according to an embodiment of the present invention. In step ST1, the called party prioritizer receives a message containing the address of a called party. The message can be a Setup message or one of the messages that follows a Setup message in an H.323 call. In step ST2, the called party prioritizer extracts the called party address from the message. In steps ST3 and ST4, the called party prioritizer determines whether the call should be given priority based on the called party address. Determining whether the call should be given priority can include searching a lookup table for the called party address to determine whether the address belongs to an important call destination, such as an emergency 911 operator. If the call is determined to be a priority call, the called party prioritizer processes the message using resources reserved for the call and bypasses the overload control routines below the called party prioritizer illustrated in Figure 4 (step ST5). The reserved resources may include processing time and/or memory accessible by the gatekeeper. If the called party prioritizer determines that the message is not addressed to a priority destination, the message is processed as normal, i.e., without reservation of resources (step ST6). Because the called party prioritizer identifies messages that relate to priority destinations for processing using reserved resources, these messages are protected from

being discarded by the remaining overload control routines illustrated in Figure 4.

Service Controller Origination Limiter

Referring back to Figure 4, another overload control routine that can be included in the service controller of a network gatekeeper is service controller origination limiter 408. Service controller origination limiter 408 limits the number of new calls accepted by the service controller in order to allow the service controller to allocate more CPU time for calls in progress. Service controller origination limiter 408 can be invoked when the processing load on the service controller reaches a threshold value greater than the threshold value for invoking zone manager admissions controller 406.

Figure 10 is a flow chart illustrating exemplary steps that may be performed by a service controller origination limiter according to an embodiment of the present invention. In step ST1, the service controller origination limiter slows the rate for accepting new TCP connections from terminals. In a multi-threaded environment, slowing the rate for accepting new TCP connections from terminals can include lowering the priority for processing server threads that accept new TCP connections from terminals. Slowing the rate for accepting new TCP connections from terminals allows the service controller to allocate more processor time for calls in progress. In step ST2, it is determined whether service controller processing load has been sufficiently reduced. If service controller processing load has been sufficiently reduced, the service controller origination limiter can process TCP server threads with normal

-22-

priority (step ST3). If service controller processing load has not been sufficiently reduced, the service controller origination limiter can slow the rate for accepting new TCP connections from gateways (step ST4).

5 Referring back to Figure 1, gateways translate between communications protocols, such as public switched telephone network protocols and H.323 protocols. Greater total network resources are expended on calls made from the PSTN network to the H.323 network than calls made
10 between terminals in the H.323 network. This is the reason that the service controller origination limiter preferably slows the rate for accepting TCP connections from terminals in the H.323 zone before slowing the rate for accepting new TCP connections from gateways. Thus,
15 the service controller origination limiter illustrates one method for efficiently regulating message processing at the service controller level.

In regulating message processing at the service controller level, it is preferable that if a Setup message
20 is discarded, the user is not billed for the call, even though bandwidth has been reserved when an ARQ message for a call is successfully processed. For example, an ARQ message can be successfully processed by the zone manager and the Setup message can be discarded by the service
25 controller. The user is preferably not billed in this situation because the call was not made. One way to ensure that the user is not billed is to provide a billing routine that checks whether a call is made, rather than checking whether a call sequence contains an ARQ and DRQ
30 to allocate and de-allocate network bandwidth.

Call Processor with In-Progress Message Favoring

-23-

Referring back to Figure 4, another level of protection that can be provided by the service controller is a call processor with in-progress message favoring 410. Call processor with in-progress message favoring 410 processes messages received by the service controller in a manner that favors calls in progress over newly originating calls. Call processor with in-progress message favoring 410 can be invoked when the service controller processing load reaches a threshold greater than the threshold required for invoking service controller origination limiter 408.

Figure 11 illustrates exemplary steps that can be performed by a call processor with in-progress message favoring. In step ST1, the call processor examines a message received by the service controller. In step ST2, the call processor determines whether or not the message is part of a newly originating call. Determining whether the message is part of a newly originating call is preferably performed in a manner that requires little processing by the service controller. For example, it is preferable that the call processor not create any new objects or parse the entire message. One method for efficiently determining whether the message is part of a newly originating call is to examine the message header and determine the message type. If the message is a Setup message, the message can be determined to be part of a newly originating call. If the message is a message that follows a Setup message, e.g., Call Proceeding, Alerting, Connect, or Release Complete, the message is determined not to be part of a newly originating call. Another method for determining whether the message is part of a newly originating call is to determine whether the connection has a call processing owner, which represents

-24-

a context established from processing of the previous messages. If the TCP connection for the message has a call processing owner, then the message represents progress work. If the TCP connection does not have a call
5 processing owner, then the message is determined to be part of a newly originating call.

If the message is determined to be part of a call in progress, the message is placed in a progress queue (step ST3). The service controller responds to messages in the
10 progress queue. The progress queue is a queue maintained by the service controller for storing call signaling messages to be processed. In step ST4, if the message is determined to be part of a newly originating call, the message is discarded. Thus, the call processor with in-
15 progress message favoring efficiently regulates calls at the service controller level.

The present invention is not limited to the steps illustrated in Figure 11 for favoring calls in progress over newly originating calls. For example, the call
20 processor with in-progress message favoring may discard the oldest origination message in the origination queue, since the oldest message is most likely to be abandoned first. The oldest origination message in the origination queue may be identified and discarded when the message is
25 extracted from the origination queue, rather than processing the message. In cases of extreme overload, the call processor with in progress message favoring may discard messages before placing the messages in the origination queue.

30 The call processor with in-progress message favoring is preferably customizable by applications. For example, an essential lines application may influence the prioritizing conditions so that messages from essential

-25-

lines will not be discarded. This customizing may be accomplished by allowing plug-ins to the call processor from applications. Another case in which it may be desirable to customize call processing is emergency 911 calls.

Abandoned Call Detector

Referring back to Figure 4, another level of protection that can be provided by service controller is an abandoned call detector 412. Abandoned call detector 412 detects calls that have been abandoned by the calling party. For example, when a calling party goes off-hook, the calling party can transmit a Setup message to the gatekeeper. If the called party does not receive a dial tone within a short period of time, the calling party can assume that his or her telephone is not functioning properly. As a result, the calling party can hang up, i.e., go on-hook, resulting in a Release Complete message being transmitted to the gatekeeper. Accordingly, if the gatekeeper receives a Setup message immediately followed by a Release Complete message pertaining to the same call, the gatekeeper can assume that the call has been abandoned. This is just one of the ways that the gatekeeper can detect abandoned calls.

Figure 12 illustrates exemplary steps that can be performed by an abandoned call detector according to an embodiment of the present invention. In step ST1, the abandoned call detector receives a first message associated with a call and stores the call reference value associated with the message. As discussed above, the call reference value is a number included in all messages associated with a call. In H.323, the first message

-26-

associated with a call can be a Setup message or an origination ARQ message. In addition to storing the call reference value, if the abandoned call detector uses the above-described method for identifying abandoned calls based on timing between messages, the abandoned call detector can also store a time value indicative of when the first message was received.

In step ST2 of Figure 12, the abandoned call detector receives the next message pertaining to the same call. Identifying the next message pertaining to the same call may include examining the call reference values of each incoming message and comparing the incoming call reference values with the call reference value of the first message. Once the next message pertaining to the same call has been received, the abandoned call detector determines whether or not the message indicates that the call has been abandoned (steps ST3 and ST4).

Determining whether the call has been abandoned can be accomplished in any number of ways. For example, the abandoned call detector can look at the time that the second message was received. If the difference between the times of receipt of the first and second messages is less than a predetermined value, the abandoned call detector can assume that the call has been abandoned. For example, referring back to Figure 3, once the client transmits a Setup message, the client has to wait for call Proceeding, Alerting, and Connect messages from the gatekeeper. Transmitting these messages to the client can require a predetermined time period, during which the calling party is supposed to remain idle. Thus, if any message is received from the client during that time period, the call can be assumed to be abandoned. Another method for determining whether the call has been abandoned

-27-

is determining whether the second message is a Release Complete message. Yet another method for determining whether the call has been abandoned can include examining the difference in time of receipt between the first and second messages and the message type of the second message. For example, if a Setup message is immediately followed by a Release Complete message pertaining to the same call, then the call is likely to have been abandoned. In step ST5, if the call has been abandoned, the abandoned call detector preferably discards all messages associated with the call. In step ST6, if the abandoned call detector determines that the call has not been abandoned, the abandoned call detector preferably places the message in a queue to be processed.

The present invention is not limited to the steps illustrated in Figure 12 for determining whether a call is abandoned. For example, the abandoned call detector can receive an origination message and remember its call reference value and time of receipt. If a predetermined time period elapses before another message is received pertaining to the same call, the abandoned call detector can assume that the call has been abandoned by the calling party. When this occurs, the abandoned call detector preferably discards the message.

25

Babbling Terminal Removal

Another overload control routine that can be included in the service controller is a routine that detects when a terminal sends a large number of messages to a service controller due to a software or hardware error in the terminal. When this occurs, the terminal is preferably removed from the I/O system of the service controller so that the service controller will not crash. For example, the service controller can maintain a count of the number of messages received from a particular terminal within a certain time period. If the number of messages and/or number of errors within a certain time period exceeds a threshold, the service controller can refuse to accept further messages from the terminal.

15 Zone Manager Fault Detection

According to another aspect of the invention, a network gateway or a terminal can include a routine for determining whether the zone manager of a gatekeeper is functioning properly. Figure 13 illustrates exemplary steps that can be performed by a zone manager fault detector according to an embodiment of the present invention. In step ST1, the zone manager fault detector transmits a ping message to the zone manager. In step ST2, the zone manager fault detector determines whether a response has been received from the zone manager. In step ST3, if a response has not been received, the zone manager fault detector determines that the protocol stack of the zone manager is down. As a result, there is no need to continue further testing and the zone manager fault detector ends.

-29-

If the zone manager fault detector receives a response to the ping message, the zone manager fault detector determines that the protocol stack of the zone manager is functioning properly (step ST4). In step ST5, the zone manager fault detector transmits a GRQ to the zone manager to determine whether the zone manager applications are functioning properly. In step ST6, the zone manager fault detector determines whether a GCF message has been received in response to the GRQ message. If a GCF message has not been received, the zone manager fault detector determines that the zone manager applications are down (step ST7). If the zone manager fault detector receives a GCF message from the zone manager, the zone manager fault detector determines that the zone manager applications are functioning properly (step ST8). In this manner, a gateway or a terminal can determine whether its zone manager is functioning properly, at both the protocol stack level and the application level.

20 Other Protocols

The overload control routines for controlling network gatekeeper message processing are not limited to the H.323 protocol. For example, some of the routines are applicable to other protocols, such as the Media Gateway Control Protocol (MGCP), for example, as described in Media Gateway Control Protocol, Version 0.1, Internet Engineering Task Force, February 21, 1999, the disclosure of which is incorporated by reference herein in its entirety. In MGCP, an entity referred to as a "call agent" or a "media gateway controller" performs similar functions to the service controller described above with regard to H.323. The MGCP call agent or media gateway

-30-

controller does not have a zone manager functionality. Accordingly, any of the routines described above with regard to overload control for the service controller are applicable to an MGCP call agent or media gateway controller. For example, referring to Figure 4, the called party prioritizer 402, the service controller origination limiter 408, the call processor with in progress message favoring 410, and the abandoned call detector 412 may each be applied to control message processing by an MGCP call agent or media gateway controller. The primary difference between applying these routines to MGCP is that the names of the messages are different from those in H.323.

Called Party Prioritizer

The steps performed by a called party prioritizer to prioritize resources for calls to priority destinations are similar to those illustrated and described for the called party prioritizer of Figure 9. However, rather than extracting the called party address from a Setup message, the MGCP called party prioritizer extracts the MGCP message containing the dialed digits. The MGCP message that contains the dialed digits is the notify message.

Call Agent or Media Gateway Controller

Origination Limiter

The steps performed by an origination limiter to limit new originations in an MGCP-compliant network are similar to those illustrated and described for the service controller originations limiter of Figure 10. However,

-31-

MGCP is UDP based, rather than TCP based. Accordingly, rather than slowing the rate for accepting TCP connections, an MGCP originations limiter may slow the rate for accepting UDP messages. That is, the new
5 originations limiter first slows the rate for accepting UDP messages from MGCP gateways and then slows the rate for accepting UDP messages from MGCP terminals or endpoints.

10

Call Processor with In-Progress Message Favoring

The steps performed by a call processor with in-progress message favoring configured to favor calls in
15 progress over newly originating calls in an MGCP-compliant network are similar to those illustrated and described for the call processor of Figure 11. In order to identify whether a message relates to a call in progress or a newly originating call, the MGCP call processor may determine
20 whether a context has been established for the call. In MGCP, receipt of the Notify message may establish a call context. Subsequent messages with the same call ID parameter as the Notify message are identified as in-progress messages. Notify messages with new call ID
25 parameters are identified as new origination messages. When an overload condition occurs, the new origination messages are preferably discarded, and the in-progress messages are retained for further processing.

Abandoned Call Detector

The steps performed by an abandoned call detector configured to detect abandoned calls in an MGCP-compliant network are similar to those illustrated and described for the abandoned call detector of Figure 12. For example, in order to detect abandoned calls in an MGCP-compliant network, the abandoned call detector may look for two messages pertaining to the same call that are closely spaced in time to each other. If the messages are close in time to each other, the first message may be assumed to be a message pertaining to an off-hook condition, and the second message may be assumed to be a message corresponding to an on-hook condition. The MGCP message that indicates to the call agent that a user is on-hook or off-hook is the Notify message. A parameter within the Notify message identifies whether the user is on-hook or off-hook. The other methods for identifying whether a call is abandoned with regard to H.323 apply equally to MGCP. For example, if a predetermined time period has passed since a message corresponding to an off-hook condition has been received, the call may be determined to be abandoned. In yet another alternative, the abandoned call detector may examine the message type to determine whether a call has been abandoned. Once a call is identified as abandoned, the abandoned call detector preferably discards the message or messages pertaining to the call.

It will be understood that various details of the invention may be changed without departing from the scope of the invention. Furthermore, the foregoing description is for the purpose of illustration only, and not for the purpose of limitation--the invention being defined by the claims.

-33-

CLAIMS

What is claimed is:

1. A gatekeeper including a zone manager and a service controller, the zone manager comprising computer executable instructions embodied in a computer readable medium for performing steps comprising:
 - (a) processing registration, admission, status (RAS) messages received by the zone manager over a network; and
 - (b) when the zone manager or the service controller enters an overloaded state, processing the RAS messages in a manner favoring calls in progress over newly originating calls, and wherein the service controller comprises computer executable instructions embodied in a computer readable medium for performing steps comprising:
 - (i) processing call signaling messages received over the network; and
 - (ii) when the service controller enters an overloaded state, processing the call signaling messages in a manner favoring calls in progress over newly originating calls.
2. The gatekeeper of claim 1, wherein processing the RAS messages includes identifying whether a calling party is a priority user and, in response to determining that the calling party is a priority user, processing messages from the calling party using resources reserved for the calling party.
3. The gatekeeper of claim 1, wherein processing the RAS messages includes preferentially discarding Admissions Request (ARQ) messages over Deallocation Request (DRQ) messages.

4. The gatekeeper of claim 1, wherein processing the RAS messages includes preferentially discarding origination Admissions Request (ARQ) messages over termination ARQ messages.

5 5. The gatekeeper of claim 1, wherein processing the call signaling messages includes identifying whether a called destination is a priority destination and, in response to determining that the destination is a priority destination, processing messages to the called destination
10 using reserved gatekeeper resources.

6. The gatekeeper of claim 1, wherein processing the call signaling messages includes slowing the rate for accepting new TCP connections from terminals in the network.

15 7. The gatekeeper of claim 6, comprising, after slowing the rate for accepting new TCP connections from terminals in the network, determining whether processing load of the service controller is reduced below a threshold value, and in response to determining that the
20 processing load is not reduced below the threshold value, slowing the rate of accepting new TCP connections from gateways in the network.

8. The gatekeeper of claim 1, wherein processing the call signaling messages includes determining whether
25 a message relates to a newly originating call or a call in progress, and in response to determining that the message relates to a newly originating call, discarding the message.

9. The gatekeeper of claim 1, comprising
30 communicating service controller processing load to the zone manager and when service controller processing load exceeds a threshold value, processing the RAS messages in

-35-

a manner that favors calls in progress over newly originating calls.

10. The gatekeeper of claim 1, wherein processing RAS messages includes processing RAS messages received
5 through an H.323 network.

11. The gatekeeper of claim 1, wherein processing call signaling messages comprises processing call signaling messages received through an H.323 network.

12. A user profile prioritizer for identifying
10 messages from priority users received by a network gatekeeper, the user profile prioritizer comprising computer executable instructions embodied in a computer readable medium for performing steps comprising:

- (a) receiving a message from a user in a network;
- 15 (b) extracting an IP address of the user from the message;
- (c) determining whether the user is a priority user; and
- (d) in response to determining that the user is a
20 priority user, processing the message using resources reserved for the user.

13. The user profile prioritizer of claim 12, wherein receiving a message from a user includes receiving
a RAS message from the user.

25 14. The user profile prioritizer of claim 13, wherein receiving a RAS message from the user includes receiving an Admissions Request message from the user.

15. The user profile prioritizer of claim 12, wherein processing the message using resources reserved
30 for the user includes bypassing overload control routines for controlling message processing during overload conditions.

16. The user profile prioritizer of claim 12, wherein receiving a message from a user includes receiving a message from an H.323 network.

17. An input handler for regulating the processing of messages received by a network gatekeeper during overload conditions, the input handler comprising computer executable instructions embodied in a computer readable medium for performing steps comprising:

(a) in response to determining that processing load of the gatekeeper exceeds a first threshold, discarding first RAS messages received by the network gatekeeper;

(b) determining whether the processing load is reduced below a second threshold smaller than the first threshold; and

(c) in response to determining that the processing load is not decreased below the second threshold, discarding second RAS messages received by the network gatekeeper.

18. The input handler of claim 17, wherein the first RAS messages are ARQ messages and the second RAS messages are DRQ messages.

19. The input handler of claim 17, wherein determining whether the processing load of the gatekeeper is reduced below the second threshold includes receiving messages indicative of the processing load of a zone manager and a service controller of the gatekeeper.

20. The input handler of claim 17, wherein discarding the first RAS messages includes discarding the first RAS messages received from an H.323 network.

21. The input handler of claim 17, wherein discarding the second RAS messages includes discarding the second RAS messages received from an H.323 network.

22. A zone manager admissions controller for regulating message processing of a gatekeeper during overload conditions, the zone manager admissions controller comprising computer executable instructions embodied in a computer readable medium for performing steps comprising:

- (a) in response to determining that processing load of the gatekeeper exceeds a first threshold, identifying and discarding origination ARQ messages received by the zone manager;
- (b) determining whether the processing load is decreased below a second threshold smaller than the first threshold; and
- (c) in response to determining that the processing load is not below the second threshold, identifying and discarding termination ARQ messages received by the zone manager.

23. The zone manager admissions controller of claim 22, wherein identifying and discarding origination ARQ messages includes:

- (a) reading a first call reference value associated with a first ARQ message received by the gatekeeper;
- (b) determining whether messages having the first call reference value have previously been received by the gatekeeper; and
- (c) in response to determining that messages having the first call reference value have been received, identifying the first message as a termination ARQ message.

24. The zone manager admissions controller of claim 23, comprising, in response to determining that messages having the first call reference value have not been

received by the gatekeeper, identifying the first message as an origination ARQ message.

25. A called party prioritizer for prioritizing processing of messages received by a network gatekeeper, the called party prioritizer comprising computer executable instructions embodied in a computer readable medium for performing steps comprising:

- (a) receiving a message from a terminal managed by the gatekeeper;
- 10 (b) extracting a called party address from the message;
- (c) determining if the message should be given priority based on the called party address; and
- (d) in response to determining that the message should be given priority, processing the message using resources reserved for the called party address.

26. The called party prioritizer of claim 25, wherein receiving a message includes receiving a Setup message from the terminal.

27. The called party prioritizer of claim 25, wherein processing the message includes bypassing overload control routines for regulating call processing during overload conditions.

28. The called party prioritizer of claim 25, wherein receiving a message includes receiving a message through an H.323 network.

29. A service controller origination limiter for limiting new call originations to a service controller of a network gatekeeper, the service controller origination limiter comprising computer executable instructions embodied in a computer readable medium for performing steps comprising:

- 5 (a) in response to determining that service controller processing load exceeds a first threshold value, slowing the rate for accepting new TCP connections from terminals managed by the gatekeeper;
- (b) determining whether gatekeeper processing load is reduced below a second threshold value smaller than the first threshold value; and
- 10 (c) in response to determining that the processing load is not decreased below the second threshold value, slowing the rate for accepting new TCP connections from gateways managed by the gatekeeper.

15 30. The service controller origination limiter of claim 29, wherein slowing the rate for accepting new TCP connections from terminals includes lowering processing priority for threads allocated for accepting the new TCP connections from the terminals.

20 31. The service controller origination limiter of claim 29, wherein slowing the rate for accepting new TCP connections from gateways includes lowering processing priority of threads allocated for accepting the new TCP connections from the gateways.

25 32. A call processor for processing calls received by a service controller of a network gatekeeper, the call processor comprising computer executable instructions embodied in a computer readable medium for performing steps comprising:

- 30 (a) examining a message received by the service controller to determine whether the message is indicative of a newly originating call or a call in progress; and

(b) in response to determining that the message is indicative of a newly originating call, discarding the message.

33. The call processor of claim 32, comprising, in
5 response to determining that the message is indicative of a call in progress, placing the message in a progress queue to be processed by the service controller.

34. The call processor of claim 32, wherein
10 examining messages received by the service controller includes examining messages received by the service controller through an H.323 network.

35. The call processor of claim 32, wherein
15 examining the message includes determining whether the message is a Setup message, and in response to determining that the message is a Setup message, identifying the message as indicative of a newly originating call.

36. An abandoned call detector for regulating
20 message processing by a network gatekeeper, the abandoned call detector comprising computer executable instructions embodied in a computer readable medium for performing steps comprising:

- (a) examining messages received by the network gatekeeper;
- (b) determining whether a calling party has
25 abandoned a call based on content and time of receipt of the messages; and
- (c) in response to determining that the calling party has abandoned the call, discarding all messages relating to the call.

30 37. The abandoned call detector of claim 36, wherein examining messages received by the network gatekeeper includes identifying a first message having a first call reference value and identifying a second message having

the first call reference value and determining whether the calling party has abandoned the call includes measuring a difference in times of receipt of the first and second messages.

5 38. The abandoned call detector of claim 36, wherein determining whether the calling party has abandoned the call includes receiving a first message and measuring elapsed time since receipt of the first message and determining that the call has been abandoned if the
10 elapsed time exceeds a threshold value.

 39. The abandoned call detector of claim 36, wherein examining messages received by the network gatekeeper includes examining messages received through an H.323 network.

15 40. A zone manager fault detector comprising computer executable instructions embodied in a computer readable medium for performing steps comprising:

- (a) transmitting a ping message to a zone manager of a network gatekeeper;
- 20 (b) determining whether a valid response to the ping message is received from the zone manager; and
- (c) in response to determining that a valid response has not been received from the zone manager, determining that the zone manager protocol stack
25 is disabled.

 41. The zone manager fault detector of claim 40, comprising, in response to receiving a valid response to the ping message, determining that the zone manager protocol stack is functioning properly.

30 42. The zone manager fault detector of claim 41, comprising:

- (a) in response to determining that the zone manager protocol stack is functioning properly, transmitting a GRQ message to the zone manager;
- (b) determining whether a GCF message is received from the zone manager; and
- (c) in response to determining that a GCF message is not received from the zone manager, determining that zone manager application programs are disabled.

43. The zone manager fault detector of claim 42, comprising, in response to receiving a GCF message from the zone manager, determining that zone manager application programs are functioning properly.

44. The zone manager fault detection of claim 40, wherein transmitting a ping message to the zone manager comprises transmitting a ping message through an H.323 network.

45. A method for regulating message processing by a network gatekeeper comprising:

- (a) receiving a plurality of messages from terminals in a zone managed by the network gatekeeper;
- (b) monitoring message processing load of the network gatekeeper;
- (c) determining whether the message processing load exceeds one or more threshold values; and
- (d) in response to determining that the message processing load exceeds one of the threshold values, preferentially processing incoming messages in a manner structured to favor calls in progress over newly originating calls.

46. The method of claim 45, wherein preferentially processing incoming messages includes preferentially

discarding incoming ARQ messages over incoming DRQ messages.

47. The method of claim 45, wherein preferentially processing incoming messages includes preferentially
5 discarding incoming origination ARQ messages over incoming termination ARQ messages.

48. The method of claim 45, wherein preferentially processing incoming messages includes slowing the rate for accepting new TCP connections from the terminals.

10 49. The method of claim 45, wherein preferentially processing incoming messages includes discarding incoming Setup messages.

50. The method of claim 45, wherein preferentially processing incoming messages includes identifying messages
15 relating to abandoned calls and discarding the messages relating to the abandoned calls.

51. The method of claim 45, wherein receiving a plurality of messages includes receiving a plurality of messages through an H.323 network.

20 52. A network gatekeeper comprising:

- (a) first processing means for processing registration, admissions, status (RAS) messages received over a network;
- (b) second processing means for processing call
25 signaling messages received over network;
- (c) means for monitoring processing loads of the first and second processing means; and
- (d) first control means associated with the first processing means for processing the RAS messages
30 in a manner favoring calls in progress over newly originating calls when the means for monitoring indicates that the processing load of the first processing means or the processing

-44-

load of the second processing means exceeds a threshold value.

53. The network gatekeeper of claim 52 comprising second control means associated with the second processing means for processing call signaling messages in a manner favoring calls in progress over newly originating calls when the means for monitoring indicates that the processing load of the second processing means exceeds a threshold value.

10 54. The network gatekeeper of claim 52 wherein the first control means comprises means for identifying whether a calling party is a priority user and, in response to determining that the calling party is a priority user, processing messages from the calling party
15 using resources reserved for the calling party.

55. The network gatekeeper of claim 52 wherein first control means comprises means for preferentially discarding Admissions Request (ARQ) messages over Deallocation Request (DRQ) messages.

20 56. The network gatekeeper of claim 52 wherein the first control means comprises means for preferentially discarding origination Admissions Request (ARQ) messages over termination ARQ messages.

25 57. The network gatekeeper of claim 53 wherein the second control means comprises means for identifying whether a called destination is a priority destination and, in response to determining that the destination is a priority destination, processing messages to the called destination using reserved gatekeeper resources.

30 58. The network gatekeeper of claim 53 wherein the second control means comprises means for slowing the rate for accepting new TCP connections from terminals in the network.

59. The network gatekeeper of claim 53 wherein the second control means comprises means for determining whether a message relates to a newly originating call or a call in progress, and in response to determining that the message relates to a newly originating call, discarding the message.

60. The network gatekeeper of claim 52 wherein the network comprises an H.323 network.

61. A method for regulating message processing by an H.323 gatekeeper or a media gateway control protocol (MGCP) call agent, the method comprising:

receiving messages over an H.323 or MGCP compliant network;

identifying messages relating to newly originating calls and messages relating to calls in progress;

processing the messages relating to calls in progress;

placing the messages relating to newly originating calls in an origination queue;

determining whether the H.323 gatekeeper or the MGCP call agent is in an overloaded condition; and

in response to determining that the H.323 gatekeeper or the MGCP call agent is in an overloaded condition, identifying the oldest message in the origination queue and discarding the message.

62. A called party prioritizer for prioritizing processing of messages received by an media gateway control protocol (MGCP) call agent, the called party prioritizer comprising computer executable instructions embodied in a computer readable medium for performing steps comprising:

(a) receiving a message from a calling party in an MGCP-compliant network;

-46-

- (b) extracting a called party address from the message;
- (c) determining whether the message should be given priority based on the called party address; and
- 5 (d) in response to determining that the message should be given priority, processing the message using resources reserved for the called party address.

63. An origination limiter for limiting new call
10 originations to a media gateway control protocol (MGCP) call agent, the origination limiter comprising computer executable instructions embodied in a computer readable medium for performing steps comprising:

- 15 (a) in response to determining that call agent processing load exceeds a first threshold value, slowing the rate for accepting UDP messages from MGCP terminals;
- (b) determining whether MGCP call agent processing load is reduced below a second threshold value
20 smaller than the first threshold value; and
- (c) in response to determining that the processing load is not decreased below the second threshold value, slowing the rate for accepting UDP messages from MGCP gateways.

25 64. An abandoned call detector for regulating message processing by a media gateway control protocol (MGCP) call agent, the abandoned call detector comprising computer executable instructions embodied in a computer readable medium for performing steps comprising:

- 30 (a) examining messages received by the MGCP call agent;

-47-

- (b) determining whether a calling party has abandoned a call based on content and time of receipt of the messages; and
- (c) in response to determining that the calling party has abandoned the call, discarding all messages relating to the call.

5

1/13

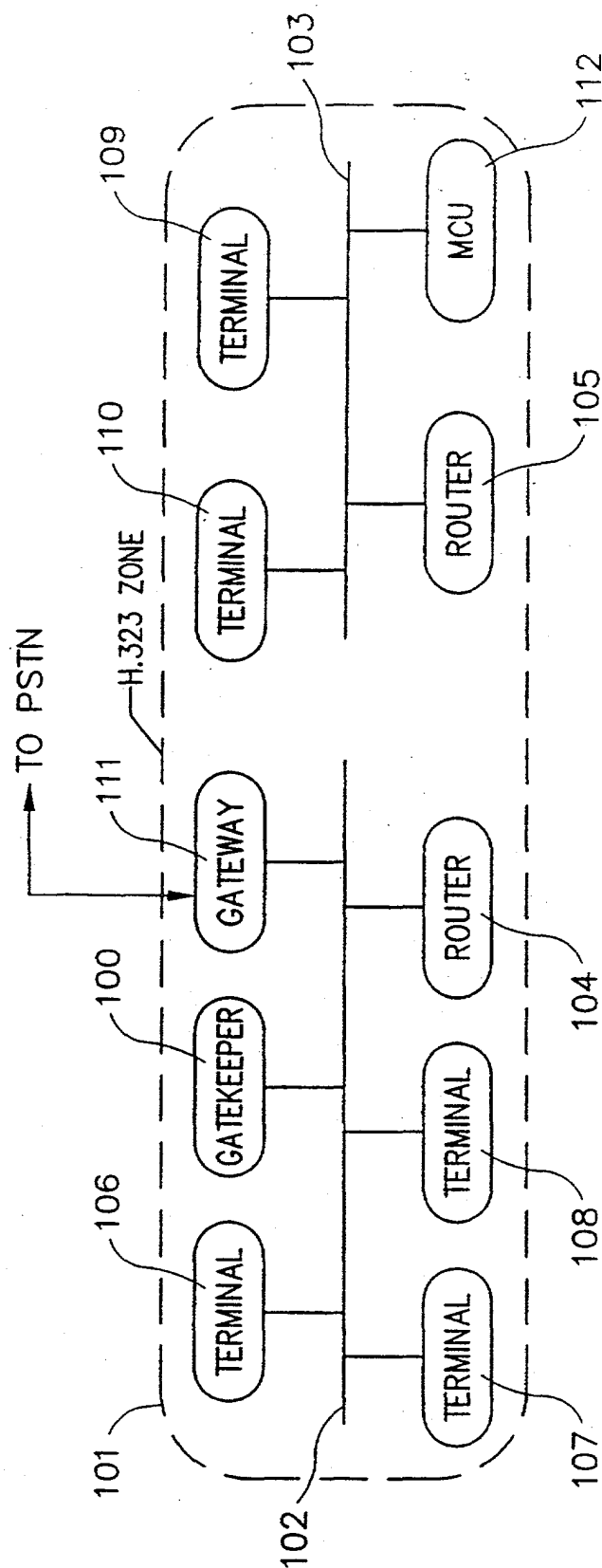


FIG.1
(PRIOR ART)

2/13

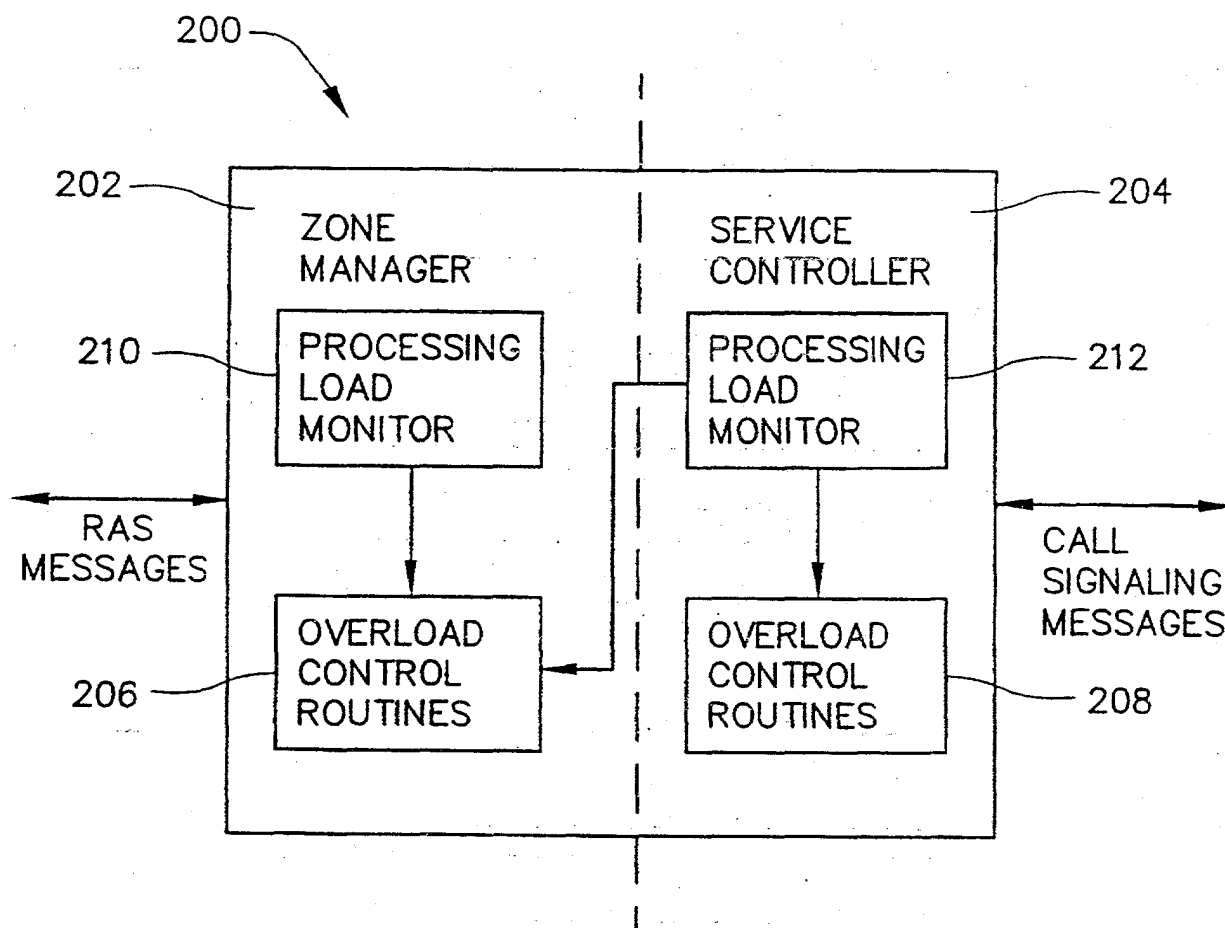


FIG.2

3/13

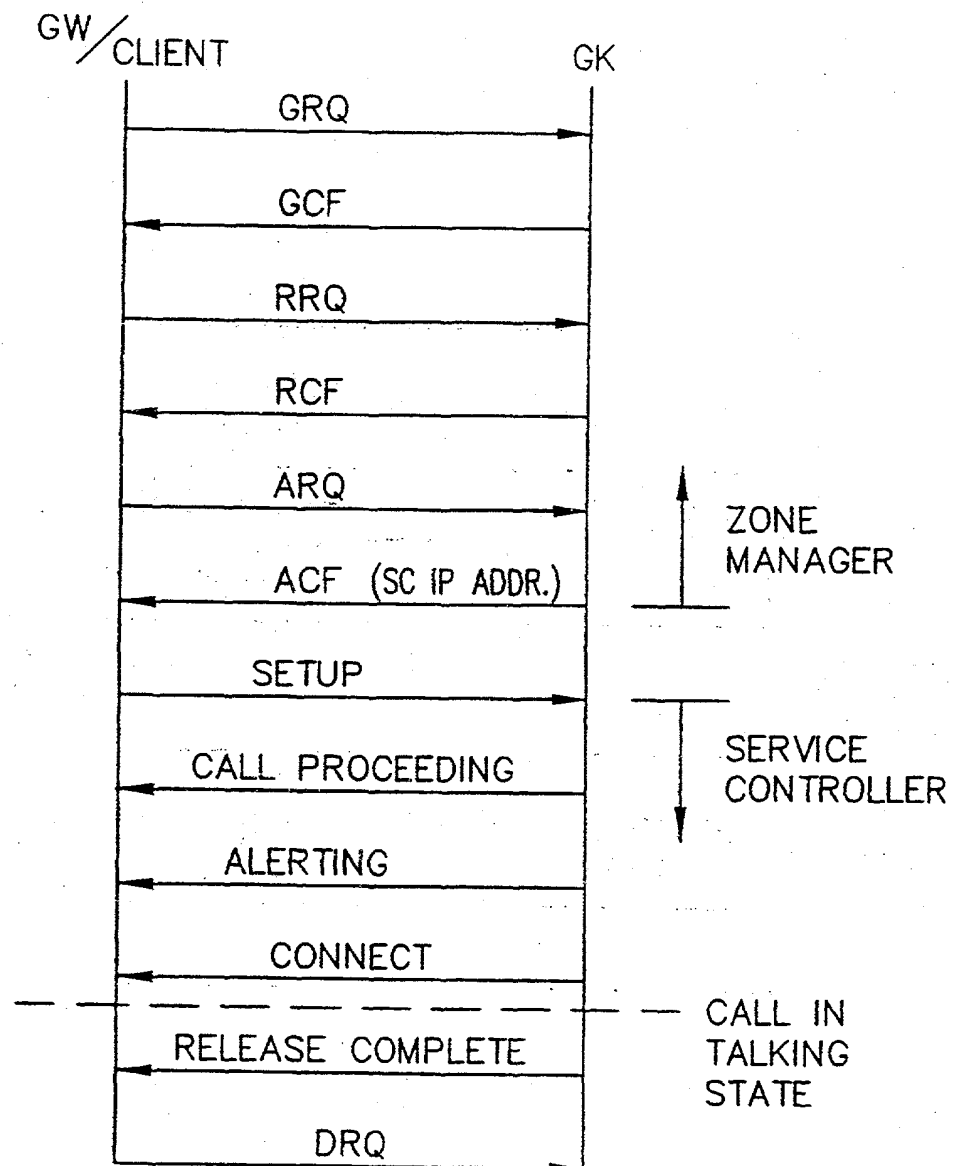
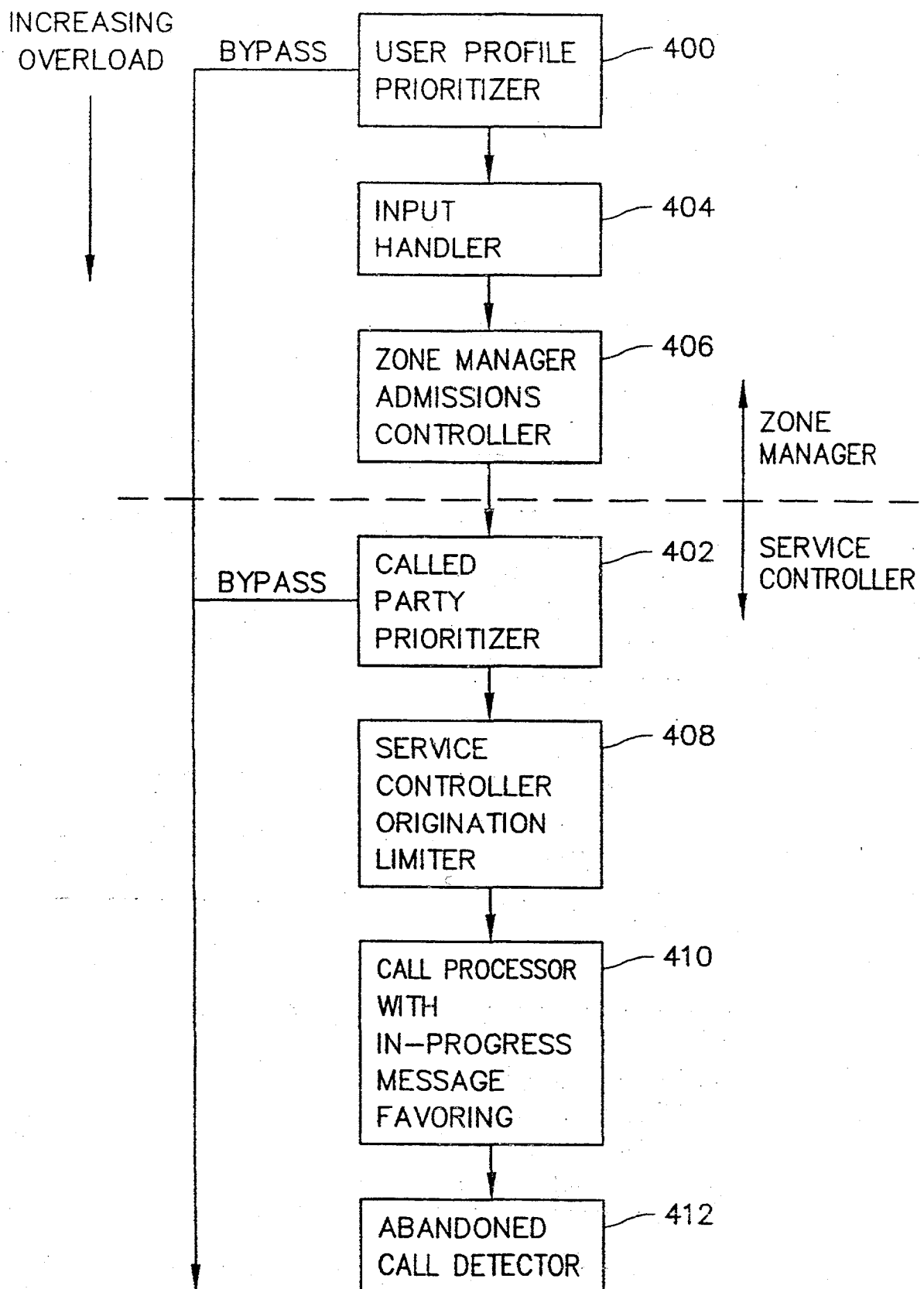
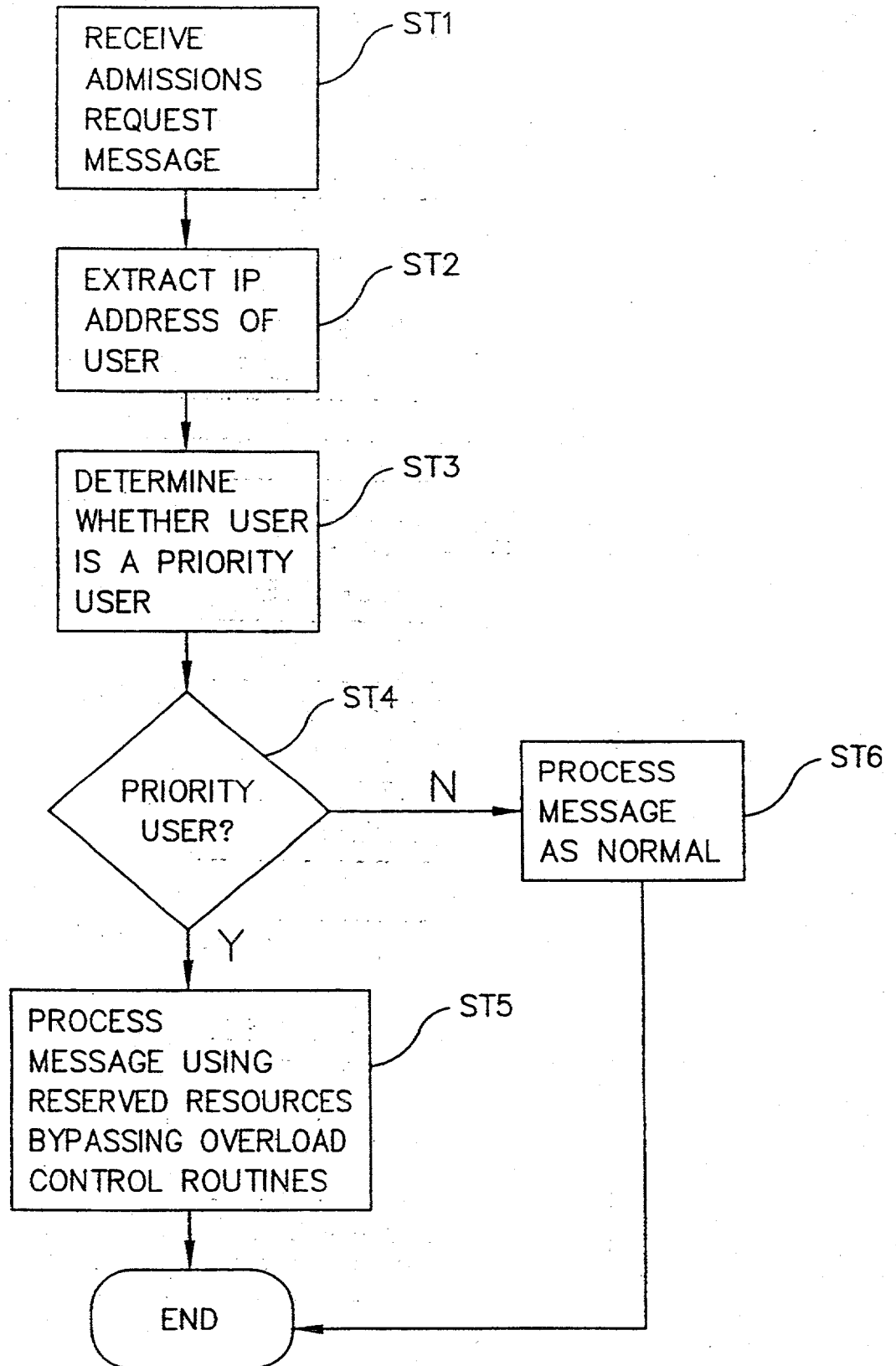


FIG.3

4/13

5/13

USER PROFILE
PRIORITIZER

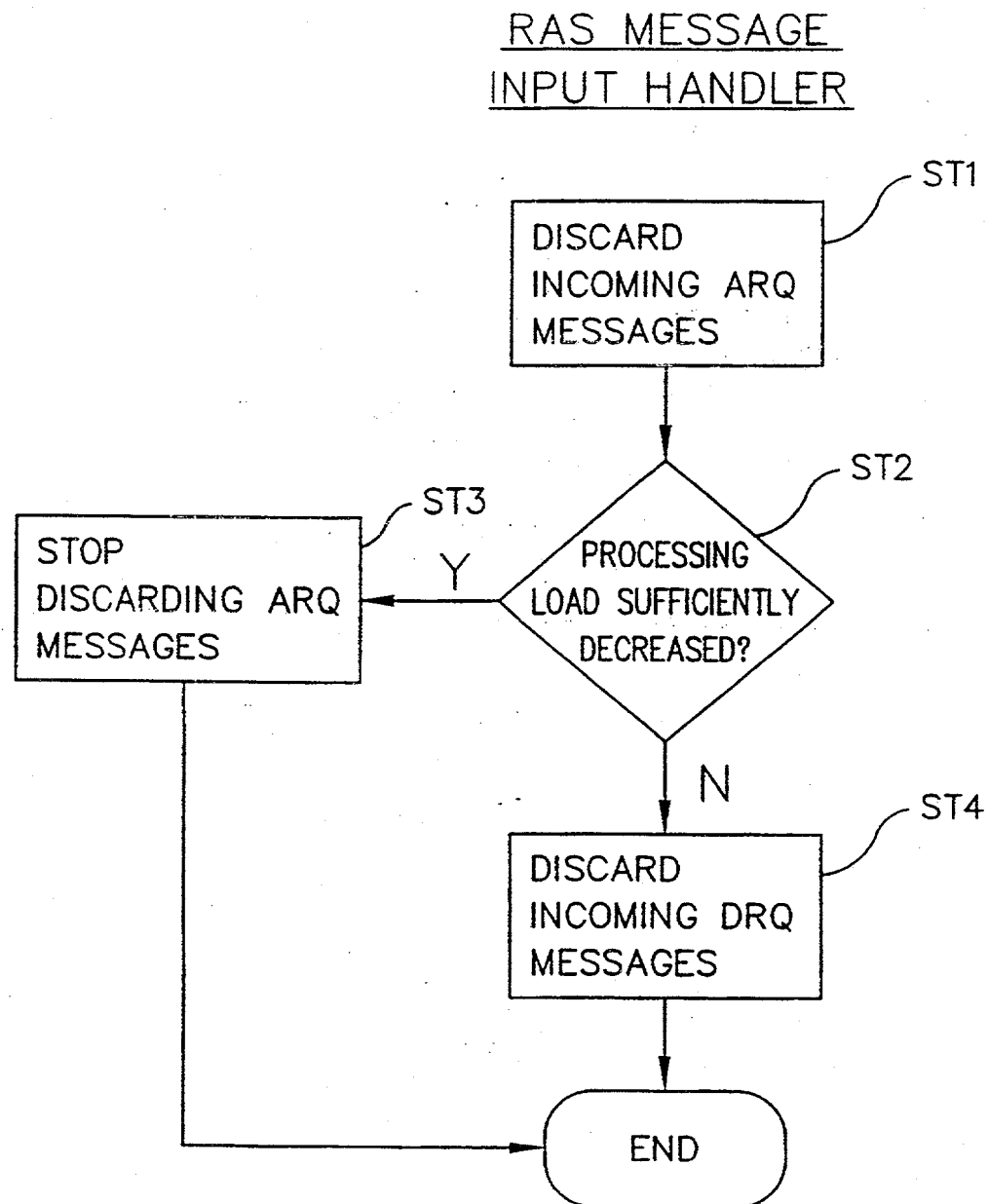
6/13

FIG.6

7/13

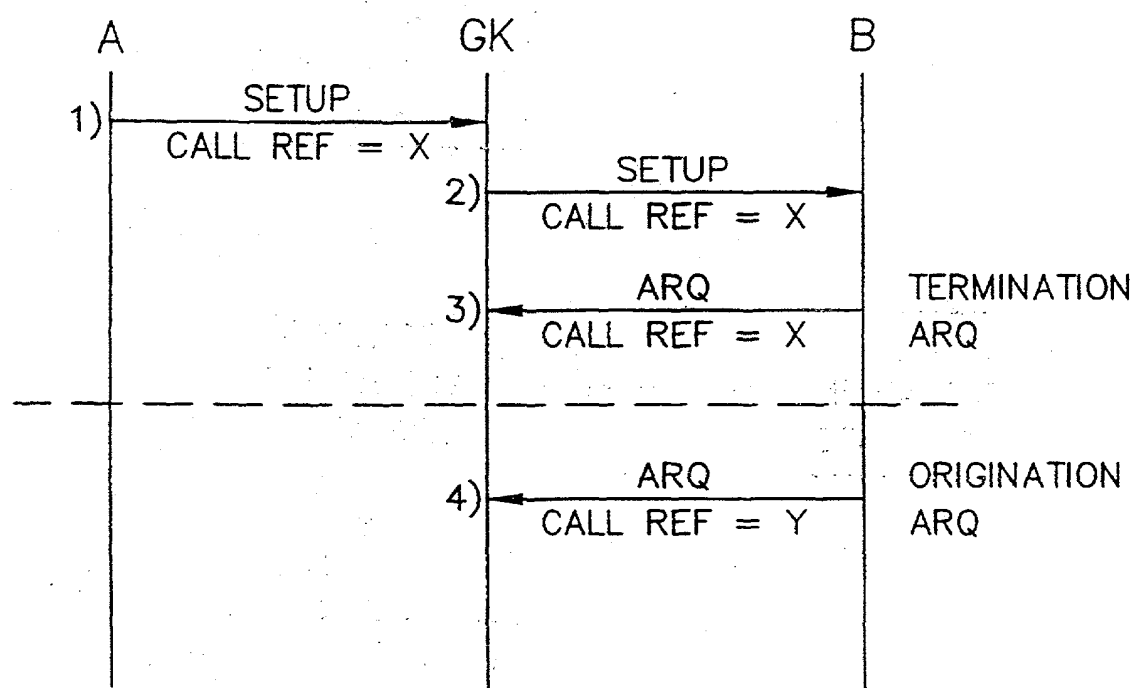


FIG.7

8/13

ZONE MANAGER ADMISSIONS
CONTROLLER

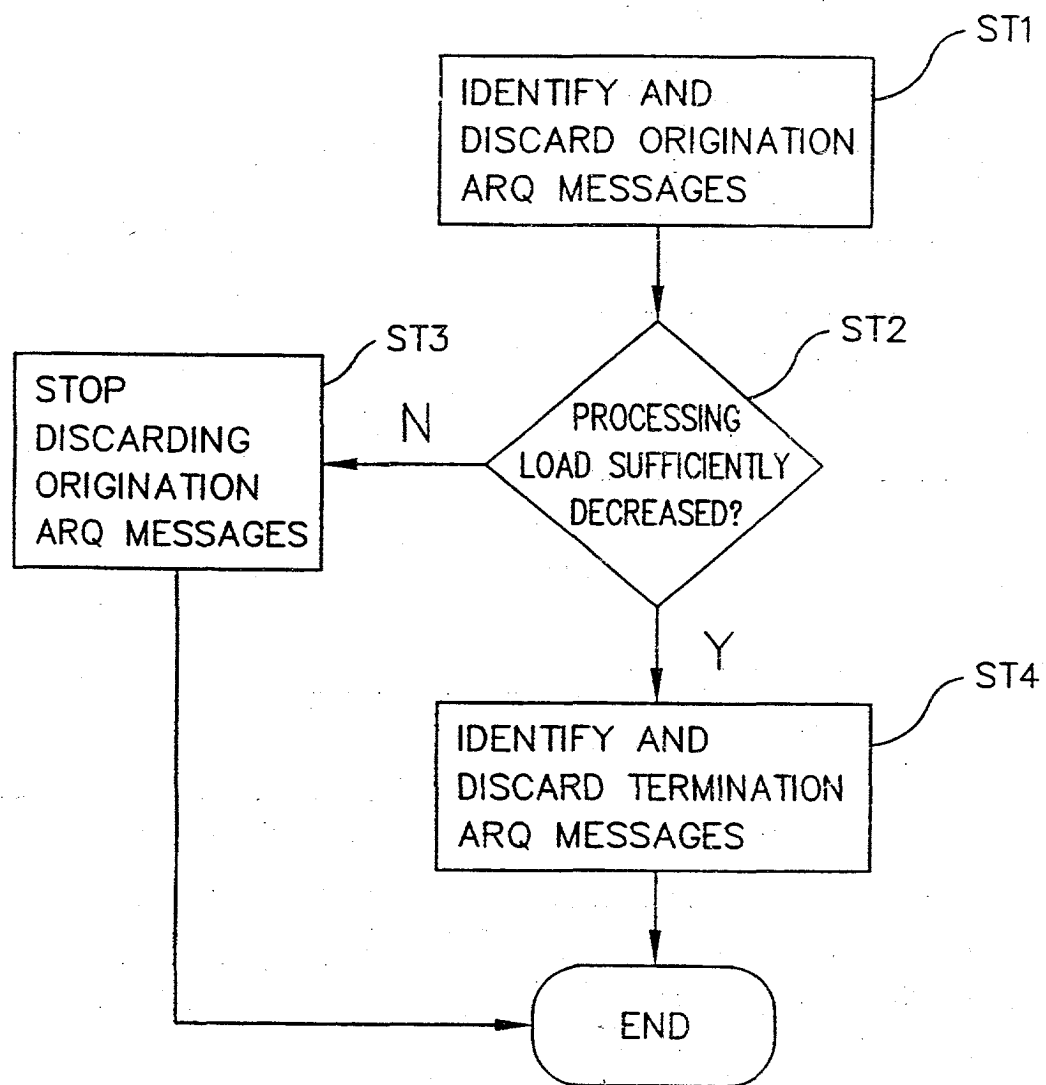


FIG.8

9/13

CALLED PARTY
PRIORITIZER

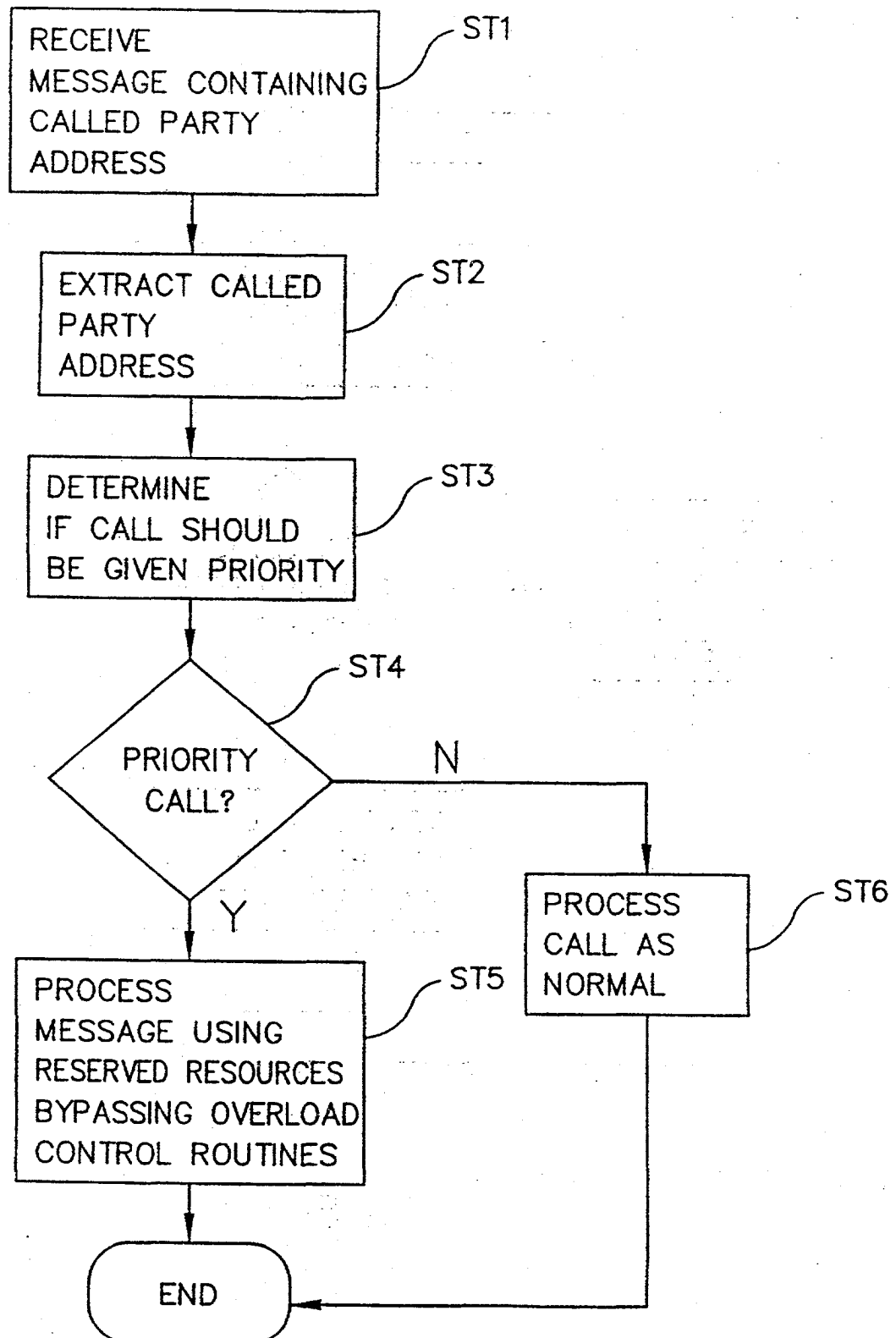


FIG. 9

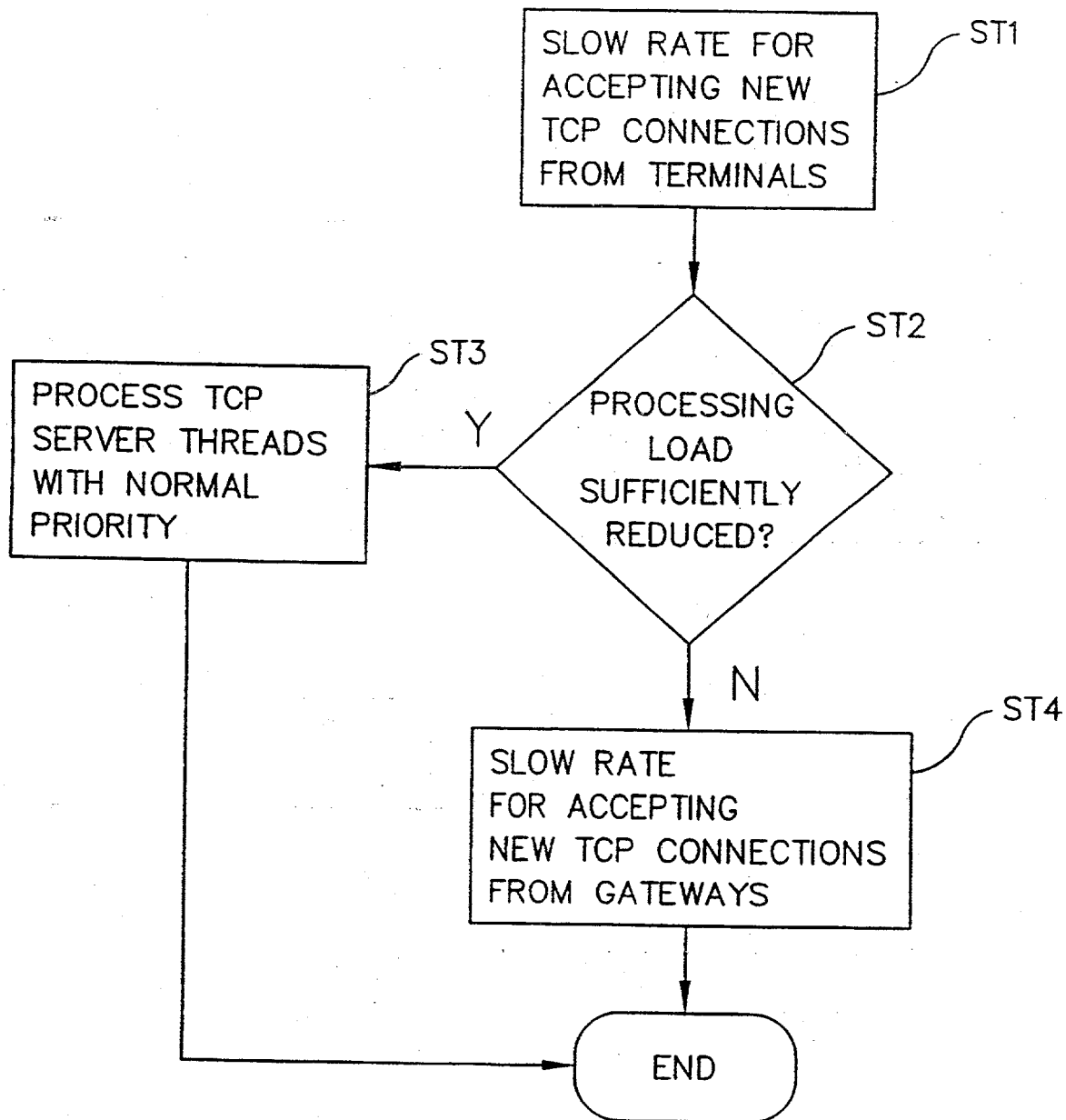
10/13SERVICE CONTROLLER ORIGINATION LIMITER

FIG.10

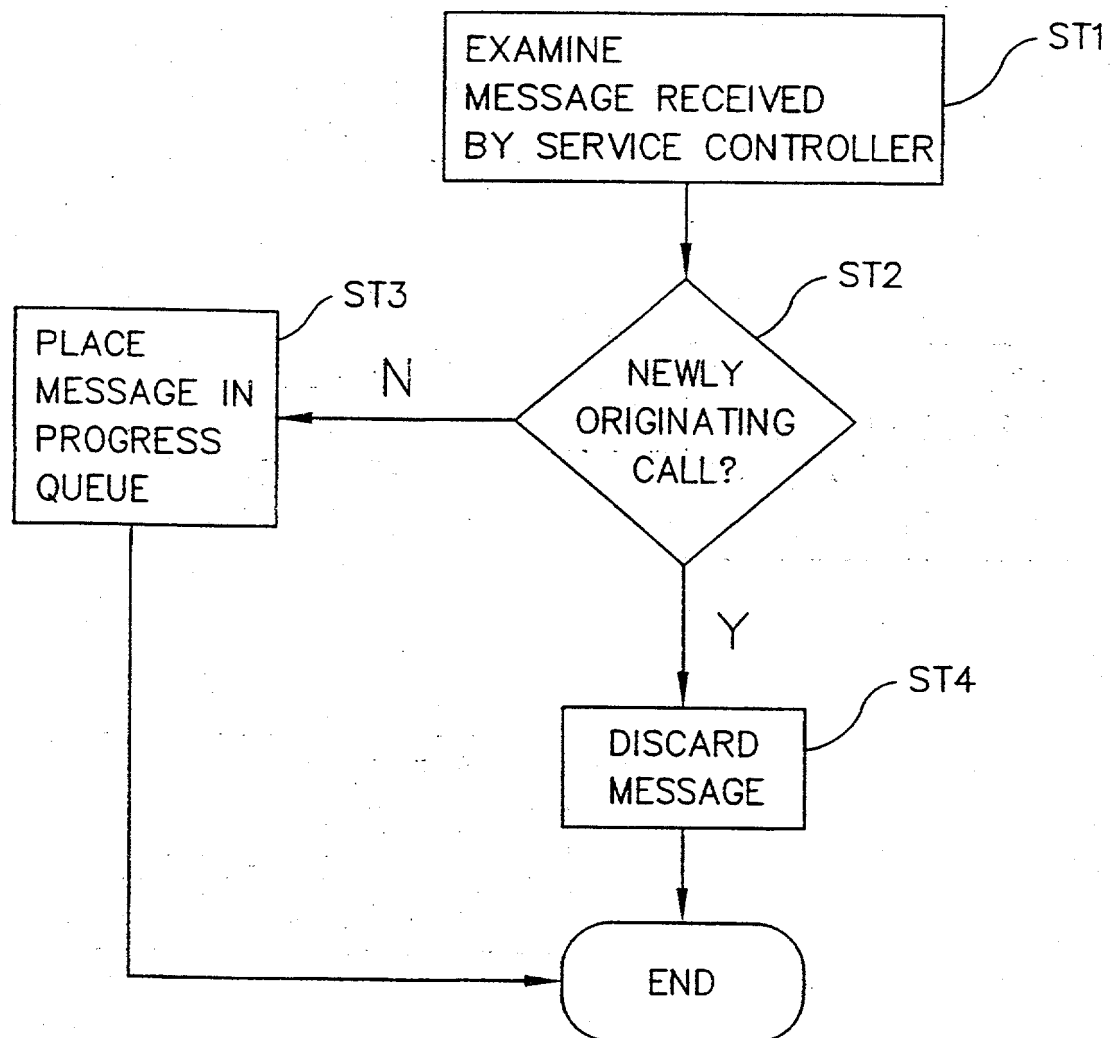
11/13CALL PROCESSOR WITH IN-PROGRESS MESSAGE FAVORING

FIG. 11

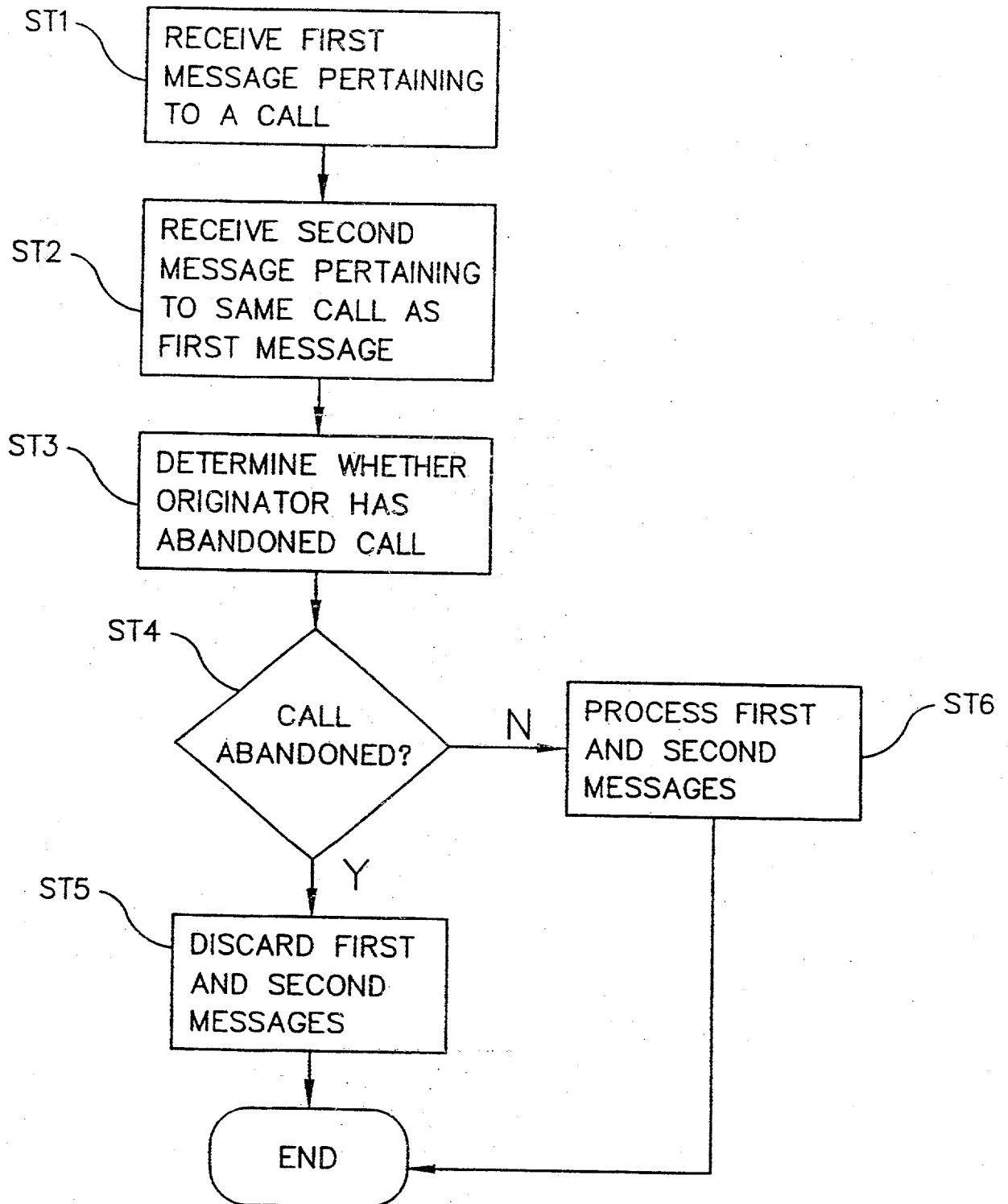
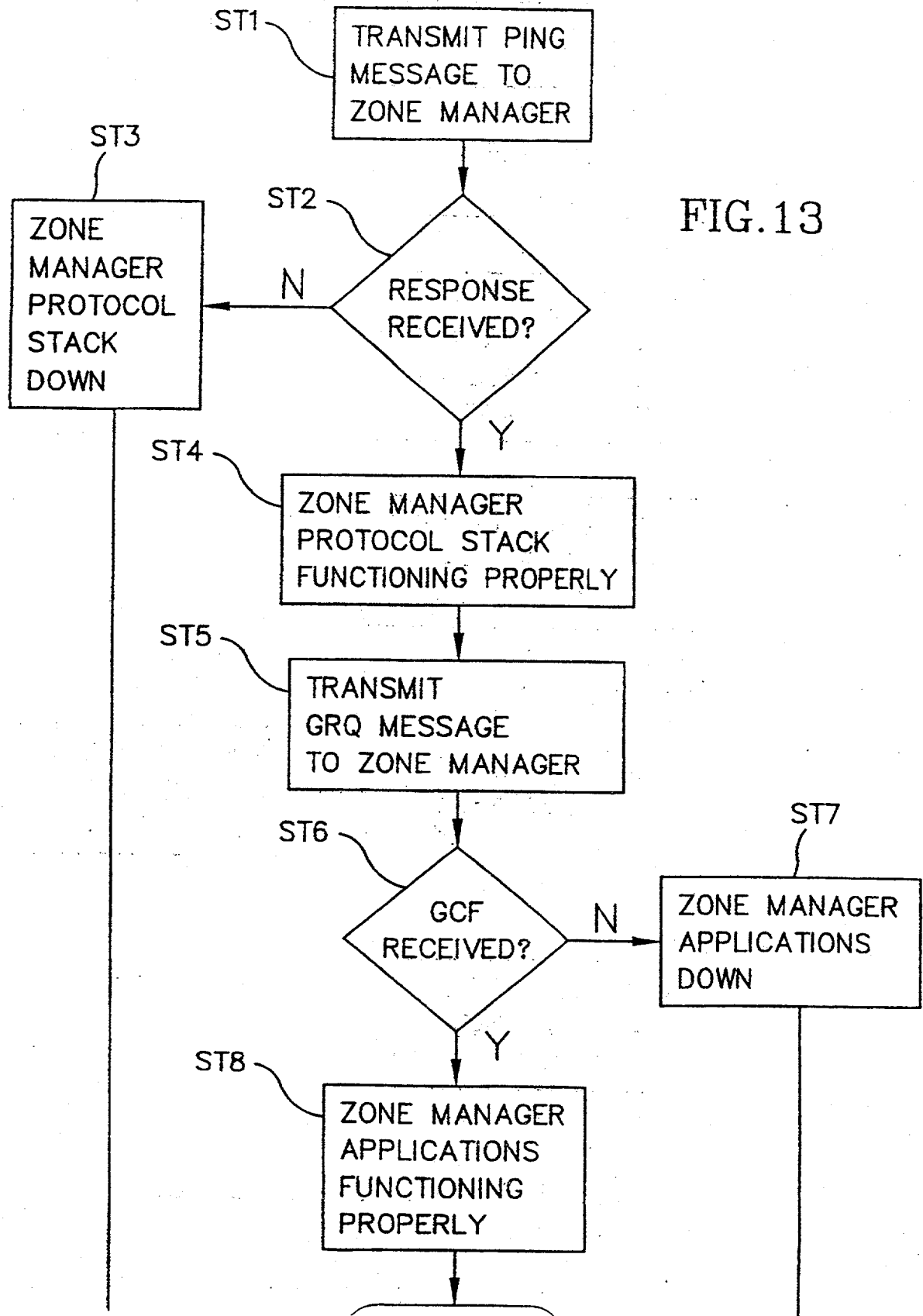
12/13ABANDONED CALL DETECTOR

FIG 12

13/13ZONE MANAGER FAULT DETECTOR

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
14 December 2000 (14.12.2000)

PCT

(10) International Publication Number
WO 00/75744 A3

(51) International Patent Classification⁷: **H04L 29/06**

(21) International Application Number: **PCT/IB00/00853**

(22) International Filing Date: **7 June 2000 (07.06.2000)**

(25) Filing Language: **English**

(26) Publication Language: **English**

(30) Priority Data:
09/327,049 **7 June 1999 (07.06.1999)** **US**

(71) Applicant: **NORTEL NETWORKS LIMITED**
[CA/CA]; World Trade Center of Montreal, 8th floor,
380 St. Antoine Street West, Montreal, Quebec H2Y 3Y4
(CA).

(72) Inventors: **LI, Xuewen**; 105 Modena Drive, Cary, NC
27513 (US). **CHRISTIE, Samuel, H., IV**; 309 Trapper's
Run Drive, Cary, NC 27513 (US).

(81) Designated States (national): **AE, AG, AL, AM, AT, AU,**
AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE,

DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU,
ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS,
LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO,
NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR,
TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): **ARIPO** patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), **Eurasian**
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), **European**
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE), **OAPI** patent (BF, BJ, CF, CG,
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

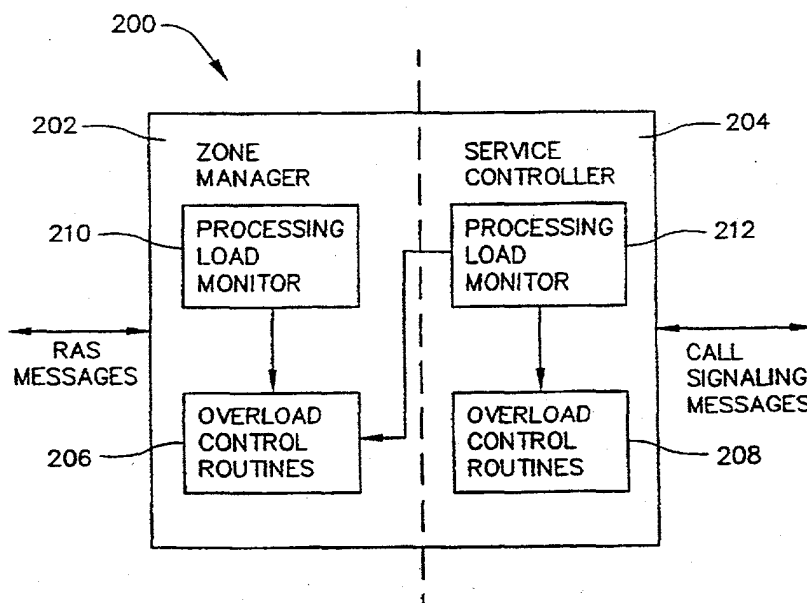
Published:

— *with international search report*

(88) Date of publication of the international search report:
23 August 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: **METHODS AND SYSTEMS FOR CONTROLLING NETWORK GATEKEEPER MESSAGE PROCESSING**



(57) Abstract: Methods and systems for controlling network gatekeeper message processing include overload control routines executed by a network gatekeeper to prevent the network gatekeeper from crashing during overload conditions. The overload control routines are executed in a predetermined order so that messages relating to calls in progress are preferentially processed over messages relating to new calls. The methods and systems also include detection and application level faults of a network gatekeeper.

O 00/75744 A3

INTERNATIONAL SEARCH REPORT

national Application No

PCT/IB 00/00853

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04L29/06

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 H04L H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, INSPEC, COMPENDEX, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>TOGA J ET AL: "ITU-T STANDARDIZATION ACTIVITIES FOR INTERACTIVE MULTIMEDIA COMMUNICATIONS ON PACKET-BASED NETWORKS: H.323 AND RELATED RECOMMENDATIONS" COMPUTER NETWORKS AND ISDN SYSTEMS,NL,NORTH HOLLAND PUBLISHING. AMSTERDAM, vol. 31, no. 3, 11 February 1999 (1999-02-11), pages 205-223, XP000700319 ISSN: 0169-7552 abstract page 206, right-hand column, line 32 -page 213, left-hand column, line 1 page 214, left-hand column, line 26 -page 219, left-hand column, line 13</p> <p style="text-align: center;">-/-</p>	<p>1,17,22, 29,32, 45,52, 61,63</p>



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents:

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

Date of the actual completion of the international search

13 December 2000

Date of mailing of the international search report

19. 03. 2001

Name and mailing address of the ISA

European Patent Office, P.O. Box 5818, Patentlaan 2

Authorized officer

INTERNATIONAL SEARCH REPORT

International Application No

PCT/IB 00/00853

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>ANQUETIL L -P ET AL: "MEDIA GATEWAY CONTROL PROTOCOL AND VOICE OVER IP GATEWAYS. MGCP AND VOIP GATEWAYS WILL OFFER SEAMLESS INTERWORKING OF NEW VOIP NETWORKS WITH TODAY'S TELEPHONE NETWORKS" ELECTRICAL COMMUNICATION, ALCATEL BRUSSELS, BE, 1 April 1999 (1999-04-01), pages 151-157, XP000830045 ISSN: 0013-4252 page 151, middle column, line 25 -page 152, left-hand column, line 32 page 153, right-hand column, line 7 -page 157, right-hand column, line 5 -----</p>	61,63

INTERNATIONAL SEARCH REPORT

international application No.
PCT/IB 00/00853

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-11, 17-24, 29-35, 45-61, 63

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-11,17-24,29-35,45-61,63

System and method comprising a gatekeeper which favors calls in progress over newly originating calls.

2. Claims: 12-16,25-28,62

System for using reserved resources for priority users.

3. Claims: 36-39,64

Abandoned call detector.

4. Claims: 40-44

Fault detector for protocol stack testing.